

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 19, 2002, 02:33:05 ; Search time 200 Seconds
(without alignments)
2438.019 Million cell updates/sec

Title: US-09-846-808-1
Perfect score: 1510
Sequence: 1 MSRRNRKFWLWMLILFTP.....ALAMLSIGYGGSGIKFIL 284

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framet_p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09846808/runat_18102002_140434_26005/app_query.fasta_1.455
-DB=N_Geneseq_032802 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosom62 -TRANS=human4.0.cdi
-LIST=45 -DOCALL=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09846808 -CGN_1.1.48 @runat_18102002_140434_26005 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	517	34.2	858	20	AAX34757	DNA encoding OMP-1
2	481.5	31.9	852	20	AAX34755	DNA encoding OMP-1
3	466	30.9	882	20	AAX34766	DNA encoding P30-6
4	455.5	30.2	900	20	AAX34768	DNA encoding P30-8
5	451	29.9	828	20	AAX34736	DNA encoding OMP-1
6	448	29.7	882	20	AAX34765	DNA encoding P30-5
7	446	29.5	894	20	AAX34749	DNA encoding OMP-1
8	432	28.6	840	20	AAX34754	DNA encoding OMP-1
9	431.5	28.6	852	20	AAX34744	DNA encoding OMP-1
10	415	27.5	891	20	AAX34767	DNA encoding P30-7
11	414.5	27.5	888	20	AAX34753	DNA encoding OMP-1
12	411.5	27.3	843	21	AAC68705	Ehrlichia chafeen
13	411.5	27.3	4683	19	AAV07179	Ehrlichia chafeen
14	411.5	27.3	4683	21	AAC68716	Ehrlichia chafeen
15	411.5	27.3	4683	22	AAV07578	DNA encoding varia
16	408.5	27.1	843	20	AAX34748	DNA encoding OMP-1
17	389	25.8	840	21	AAV01294	Ehrlichia canis im
18	389	25.8	2037	21	AAV01295	Ehrlichia canis im
19	387.5	25.7	837	20	AAX34747	DNA encoding OMP-1
20	387.5	25.7	837	21	AAC68704	Ehrlichia chafeen
21	386	25.6	843	20	AAX34762	DNA encoding P30-2
22	386	25.6	867	20	AAX34759	DNA encoding P30 p
23	383.5	25.4	864	19	AAV07176	Cowdria ruminantlu
24	383.5	25.4	864	21	AAC68699	DNA encoding major
25	383.5	25.4	864	22	AAV07575	DNA encoding major
26	380	25.2	861	20	AAX34746	DNA encoding OMP-1
27	380	25.2	861	21	AAC68703	Ehrlichia chafeen
28	379.5	25.1	845	20	AAX34743	DNA encoding OMP-1
29	377	25.0	817	20	AAX34758	DNA encoding OMP-1
30	375.5	24.9	756	20	AAX34742	DNA encoding p28 p
31	371.5	24.6	830	21	AAC68706	Ehrlichia chafeen
32	371.5	24.6	843	20	AAX34745	DNA encoding OMP-1
33	370	24.5	842	19	AAV07177	Ehrlichia chafeen
34	370	24.5	842	21	AAC68700	Ehrlichia chafeen
35	370	24.5	842	22	AAV07576	DNA encoding major
36	363	24.0	924	20	AAX34761	DNA encoding P30-1
37	363	24.0	1607	21	AAV01292	Ehrlichia canis im
38	342.5	22.7	1570	19	AAV07180	Ehrlichia canis VS
39	342.5	22.7	1570	22	AAV07583	DNA encoding varia
40	341.5	22.6	831	20	AAX34764	DNA encoding P30-4
41	341.5	22.6	864	21	AAC68707	Ehrlichia canis VS
42	339.5	22.5	1570	21	AAC68717	Ehrlichia canis 28
43	335.5	22.1	864	20	AAX34760	DNA encoding P30a
44	325.5	21.6	852	20	AAX34763	DNA encoding P30-3
45	324.5	21.5	849	21	AAV01293	Ehrlichia canis im

ALIGNMENTS

RESULT 1
AAX34757
ID AAX34757 standard; DNA; 858 BP.
XX
AC AAX34757;
XX
XX 05-JUL-1999 (first entry)
DT
XX DNA encoding OMP-1Y protein.
DE
XX Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW detection; dog; ss.
XX
XX Ehrlichia chafeensis.
OS
XX WO9913720-A1.
XX
XX 25-MAR-1999.
PD

```

XX 18-SEP-1998; 98WO-US19600.
XX
XX 19-SEP-1997; 97US-0059353.
XX
XX (OHIS ) UNIV OHIO STATE.
XX
XX Ohashi N, Rikihisa Y;
XX
XX WPI; 1999-254290/21.
XX P-PSDB; AAY06957.
XX
XX Novel outer membrane proteins from Ehrlichia chaffeensis and
XX Ehrlichia canis
XX
XX Disclosure; Fig 17A; 55pp; English.
XX
XX The invention provides isolated outer membrane proteins (OMP) from
XX Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
XX of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
XX in AAY06943-958. The E. canis proteins form part of the P30 family and
XX consist of proteins shown in AAY06959-970. The proteins and genes are
XX used to detect E. chaffeensis in patients and E. canis in dogs.
XX
XX Sequence 858 BP; 293 A; 135 C; 158 G; 272 T; 0 other;

Alignment Scores:
Pred. No.: 8,84e-49 Length: 858
Score: 517.00 Matches: 109
Percent Similarity: 59.59% Conservative: 65
Best Local Similarity: 37.33% Mismatches: 96
Query Match: 34.24% Indels: 22
DB: 20 Gaps: 8

US-09-846-808-1 (1-284) x AAX34757 (1-858)

QY 5 SerAsnArgLys-----PheValLeuTrpValMetLeu-----ileLeuPheThr 19
Db 4 AATAATAGAAAAAGTTTATATAATAGGTGCATCTACTACGAAGCTTATTATTACCA 63
QY 20 ProHisIleSerLeuAlaSerValLeuAsnAspHisAsn-----SerMetTyr 35
Db 64 TCTGAGCCCTCTTCTACAGAAATGTAAGTAACCATCTATTATTTAAACCTAGGTTATAT 123
QY 36 ValcylileGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSer 55
Db 124 ATCAGTGGACAATATAGACCAGAGCTTCTCATTTTACGAAATTTTCAGTCAAGAAACC 183
QY 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThr 75
Db 184 AACTACAATACTACTCAACTAGTTGGCTTAAAAAGGACATCATGTCATAGGGAACAGT 243
QY 76 GlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArg 95
Db 244 AATATCACACCTACACAAATTTCAACTTTCCTTACATTGCAGAAATTTCAAGACAATGCC 303
QY 96 LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGlu 115
Db 304 ATAGTTTTCAGTGGGCAATTTGATCTGATTCGAGAAATTTTGAATTTGAAGTAGAG 363
QY 116 LeuSerTyrGluThrPheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLys 135
Db 364 GCTTCTTATGAACAATTTGATGTTAAAAATCCAGAGGAGTCTCTACAGCGCATACAGG 423
QY 136 HisPheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsn----- 153
Db 424 TATTTTGCACCTAGCAGCTGCTATGGATGGCACTAAATAAATCTAGTCTCTGATGACACAAGA 483
QY 154 LysTyrValThrLeuLeuAsnAsnGlyIleSerLeuThrSerAlaLeuIleValCys 173
Db 484 AATTCACCTCATGAGAAATAGCGGGTATCAATTTTCATCAGTATGATGAATGGGTGT 543
QY 174 TyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPhe 191

```

```

Db 544 TACAATTTTACATAGATATATACCA-----GTAGTACCGTATGTATGCGCAGGAATA 597
QY 192 GlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLys 211
Db 598 GGAGGAGATTTTCAATAGAGTTTAAATGATTTACATGTTAGTTTGTCTATCAACGGAAG 657
QY 212 LeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHis 231
Db 658 GTAGGTATTAGTTATTCTATATCCCTGAAGTAAGTTTATTCTTAACGGATATTACCAT 717
QY 232 GlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSer 251
Db 718 AAAGTAACAGCTAACAGATTTAAAACTTACACGTTCAACACGTAAGTGATTTA---AGT 774
QY 252 ProThrProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIle 271
Db 775 GAGCCTCCTAAG-----TTTCACATCTGCGAGTTGCTACACTCAATGTT 816
QY 272 GlyTyrTyrGlyGlySerIleGlyIleLysPheIle 283
Db 817 GGGTACTTGGTGGCGAAATTTGGAGTAAGATTTATA 852

RESULT 2
AAX34755
ID AAX34755 standard; DNA; 852 BP.
XX
XX AAX34755;
XX
XX 05-JUL-1999 (first entry)
XX
XX DNA encoding OMP-1W protein.
XX
XX KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
XX KW detection; dog; ss.
XX
XX OS Ehrlichia chaffeensis.
XX
XX PN WO9913720-A1.
XX
XX PD 25-MAR-1999.
XX
XX PF 18-SEP-1998; 98WO-US19600.
XX
XX PR 19-SEP-1997; 97US-0059353.
XX
XX PA (OHIS ) UNIV OHIO STATE.
XX
XX PI Ohashi N, Rikihisa Y;
XX
XX DR WPI; 1999-254290/21.
XX
XX DR P-PSDB; AAY06955.
XX
XX Novel outer membrane proteins from Ehrlichia chaffeensis and
XX Ehrlichia canis
XX
XX PS Disclosure; Fig 15A; 55pp; English.
XX
XX CC The invention provides isolated outer membrane proteins (OMP) from
XX Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
XX of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
XX in AAY06943-958. The E. canis proteins form part of the P30 family and
XX consist of proteins shown in AAY06959-970. The proteins and genes are
XX used to detect E. chaffeensis in patients and E. canis in dogs.
XX
XX SQ Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 other;

Alignment Scores:
Pred. No.: 9.01e-45 Length: 852
Score: 481.50 Matches: 106
Percent Similarity: 55.60% Conservative: 48
Best Local Similarity: 38.27% Mismatches: 108
Query Match: 31.89% Indels: 15

```


AC AAX34756;
 XX 05-JUL-1999 (first entry)
 XX DNA encoding OMP-1X protein.
 XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX Ehrlichia chaffeensis.
 OS
 XX
 XX W09913720-A1.
 PN
 XX
 XX 25-MAR-1999.
 PD
 XX 18-SEP-1998; 98WO-US19600.
 PF
 XX 19-SEP-1997; 97US-0059353.
 PR
 XX (OHIS) UNIV OHIO STATE.
 PA
 XX Ohashi N, Rikihisa Y;
 PI
 XX WPI; 1999-254290/21.
 DR
 DR P-PSDB; AAY06956.
 XX
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 PS Disclosure; Fig 16A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 XX Sequence 828 BP; 281 A; 131 C; 136 G; 280 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 2.43e-41 Length: 828
 Score: 451.00 Matches: 101
 Percent Similarity: 53.76% Conservative: 49
 Best Local Similarity: 36.20% Mismatches: 99
 Query Match: 29.87% Indels: 30
 DB: 20 Gaps: 6

US-09-846-808-1 (1-284) x AAX34756 (1-828)

QY 14 MetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHis----- 31
 DB 37 CTATTTCATATGTTGTTACTTAACATATCTTTCCAGAAACATATTAAACAATACACTGAT 96
 QY 32 -----AsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB 97 AACATTTCTGGGTATATATAAGTGGCAATATAACACAGGGATTTCTCATTTCCACAAA 156
 QY 50 LeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheClyLeuLysLysAspLeu 69
 DB 157 TTTTCAGTCAAGAAATCTATAATGATAACATTCAACTAATTTGGGTAAAGACACACGCA 216
 QY 70 LeuAsnAspLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnPro 89
 DB 217 ATTTCTACTAGTACCTTAAATATT-----AATACAGATTTTAATATCCCTATAAAGTA 270
 QY 90 TyrTyrGluAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsn 109
 DB 271 ACATTTCAAAATACATTAACAGCTTTAGTGGAGCTATTGGTTATCTCATCCACAGGG 330
 QY 110 PheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsnGlyTyrLys 129
 DB 331 GCAAGATTTGAGCTTGAAGGTTCTTATGAAGAATTGTGATGCACATCCTGGAGACTGC 390

QY 130 ArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsn 148
 DB 391 TTAATAAAGATACCTATAGATATTTTCGTTTAGTAGAACCACATCAGGTTCTAGC--- 447
 QY 149 AsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAla 168
 DB 448 ---CCTACCTCAACAACATACTGTTATGAGAAATCATGCTGTTTCCATTACTTCTGTT 504
 QY 169 LeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThrTyrSerCys 188
 DB 505 ATATTTAATGGCTGTTATGACATCTTTTAAAGGATTTAGAAATATCACCTTAATGATGT 564
 QY 189 LeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyr 208
 DB 565 GTTGGTGTAGTGGAGATTTTATAGAATTTTTCGACCATTTACACATTAATAATTAGCATAC 624
 QY 209 GlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGly 228
 DB 625 CAAGGCAAGTTAGTATCAATTATCACTTATCGACTCAAGCAAGCGTATTATTATGATGGA 684
 QY 229 TyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrProCysAsp 248
 DB 685 TATTATCAATAAGTTATAGGAATCAATTCACAATCTAAATGTTCAA----- 732
 QY 249 TyrProSerProThrProProAsnSerLysProHisValHisThrThr----- 264
 DB 733 -----CACGTGGCTAGTACAGATTTTGGACCT 759
 QY 265 -----AlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLys 281
 DB 760 GTATACGCCAGTACCCACTTAACATTTGTTATTTGGTGGTGAATCGGAATTAGA 816

RESULT 6
 AAX34765
 ID AAX34765 standard; DNA; 882 BP.
 XX
 AC AAX34765;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE DNA encoding P30-5 protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX
 OS Ehrlichia canis.
 XX
 PN W09913720-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19600.
 PR
 XX 19-SEP-1997; 97US-0059353.
 XX (OHIS) UNIV OHIO STATE.
 PA
 XX Ohashi N, Rikihisa Y;
 PI
 XX WPI; 1999-254290/21.
 DR
 DR P-PSDB; AAY06965.
 XX
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 PS Disclosure; Fig 25A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are

[illegible]

Db 643 GGTATTAGTTACTTCTTCCTAAGATTAAATGATTTTGTGGTGGTACTATCATAGA 702
 Qy 233 LeuPheGlyLysLysPheGluLysIleProValAsnTyrProCys-----AspTyr 249
 Db 703 GTATAGGGAATAAATTTAAAAATTTAAATGTTAACCATGTTGTACACTTGATGAATTT 762
 Qy 250 ProSerProThrProProAsnSerLysProHisValHisThrAlaLeuAlaMetLeu 269
 Db 763 CCTAAGCA-----ACTTCTGCAGTAGCTACACTT 792
 Qy 270 SerIleGlyTyrTyrGlyGlySerIleGlyLysPhe 282
 Db 793 AATGTTGTTATTGTTGGTGAAGCTGGAGTAAAGTTT 831
 RESULT 9
 AAX34744
 ID AAX34744 standard; DNA; 852 BP.
 XX
 AC AAX34744;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE DNA encoding OMP-1B protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX
 OS Ehrlichia chaffeensis.
 XX
 FN WO9913720-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19600.
 XX
 PR 19-SEP-1997; 97US-0059353.
 XX
 PA (OHIS) UNIV OHIO STATE.
 XX
 PI Ohashi N, Rikihisa Y;
 XX
 DR WP1; 1999-254290/21.
 DR P-PSDB; AAY06944.
 XX
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 PS Disclosure; Fig 4A; 55pp; English.
 XX
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 SQ Sequence 852 BP; 301 A; 155 C; 157 G; 239 T; 0 other;
 Alignment Scores:
 Pred. No.: 4.05e-39 Length: 852
 Score: 431.50 Matches: 105
 Percent Similarity: 51.90% Conservative: 45
 Best Local Similarity: 36.33% Mismatches: 100
 Query Match: 28.58% Indels: 39
 DB: 20 Gaps: 12
 US-09-846-808-1 (1-284) x AAX34744 (1-852)
 Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAla----- 25
 Db 37 ATTTCATTAAATGAATCTTACCTTACCAATCTTTTCAGATCTCTGAACCTCAAAATGAT 96
 Qy 26 SerValLeuAsnAsp---HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArg 44

Db 97 ACAGGAATCAACGACAGCAGAGAGGCTTCATACATTAGTGTAAAGTATAATCCAGACATA 156
 Qy 45 GlnHisLeuSerLysLeuLeuIleLysGluSerAlaIleAsn-----Thr 59
 Db 157 TCACACTTCAGAAATTTCTCAGCTGAAGAAGCTCCCATCAATGGAATACTTCTCATCAT 216
 Qy 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
 Db 217 AAAAAGGTTTTCGGGCTGAAAAAGAC-----GGAGATATAGACACAA 258
 Qy 80 AsnThrAsnPheAsnIleLysTyrAsnPro-----TyrTyrGluAsnAsnArgLeu--- 96
 Db 259 TCTCGAATTTTAAAC---AGGACAGATCCAGCCCTCGAGTTTCAGATAAACCTAATATCA 315
 Qy 97 GlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeu 116
 Db 316 GGATTCTCAGGAAGTATTGGTTATGCTATGGAT---GGGCCAAGAATAAGAACTTGAAGCT 372
 Qy 117 SerTyrGluThrPheHisIleLys-----AsnAsnGlyTyrLysArgIleAspCys 133
 Db 373 GCATACCAAAATTTGATGCAAAAATCTGACAACAATGACACTAATAGCGGTGACTAC 432
 Qy 134 GluLysHisPheAlaLeuAlaLysGluIleSerGlyLysSerAsnAsnProAlaAsn 153
 Db 433 TATAAATCTTTGGACTATCTCGTCAA-----GACCAATAGCAGATAAG 477
 Qy 154 LysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCys 173
 Db 478 AAATATGTTGCTTAAAAATGAAGCATCACTTTTATGTCATTATGTTAACAACCTTGC 537
 Qy 174 TyrAspValAspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyVal 193
 Db 538 TATGACATTACAGCTGAAGGAGTACCTTTCATCCGTATGTCATGCGAGGTGTAGAGCA 597
 Qy 194 AspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGly 213
 Db 598 GACCTTATAAAGCTATTTAAGGATTTTAAATTTAAATTTCTATCAACCAAGGAAATAGGT 657
 Qy 214 AlaSerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeu 233
 Db 658 ATTAGCTATCCAATCACACCAGAGTTTCCGCTTTTATTGGAGGATACCTACCACGAGTT 717
 Qy 234 PheGlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThr 253
 Db 718 ATAGAAATAATTTTAAACAAATACCTGTA-----ATAACACCTGTA 759
 Qy 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
 Db 760 GTATTAGAAGGAGCTCTCTCAAAACC---ACATCTGCGCTAGTAACTATTGACACTGGATAC 816
 Qy 274 TyrGlyGlySerIleGlyIleLysPhe 282
 Db 817 TTGCGCGAGAAGTTGGAGTAAGGTTTC 843
 RESULT 10
 AAX34767
 ID AAX34767 standard; DNA; 891 BP.
 XX
 AC AAX34767;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE DNA encoding P30-7 protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX
 OS Ehrlichia canis.
 XX
 FN WO9913720-A1.
 XX
 PD 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19600.
 XX 19-SEP-1997; 97US-0059353.
 XX (OHIS) UNIV OHIO STATE.
 PA Ohashi N, Rikihisa Y;
 PI WPI; 1999-254290/21.
 DR P-PSDB; AAY06957.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 27A; 55pp; English.
 PS The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the p30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 891 BP; 327 A; 135 C; 130 G; 299 T; 0 other;
 SQ Alignment Scores:
 Pred. No.: 3.16e-37 Length: 891
 Score: 415.00 Matches: 99
 Percent Similarity: 49.68 Conservative: 55
 Best Local Similarity: 31.94 Mismatches: 114
 Query Match: 27.48 Indels: 42
 DB: 20 Gaps: 8
 US-09-846-808-1 (1-284) x AAX34767 (1-891)
 Qy 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIleLeuPheThr--- 19
 Db 1 ATGGGAAATCTATGAATAAATAAGTCAATCTTAATAAGATTATATATTTTAACATGC 60
 Qy 20 -----ProHisIleSerLeuAlaSerValLeuAsnAsp---HisAsnSerMet 34
 Db 61 ATGCTGTCATTACCTTAATATATCTCTTCAAAAGTAATAACGAAACATCTCGTTG 120
 Qy 35 TyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuIleLysGlu 54
 Db 121 TATATTAGCGGCATACAAACCCAGTCTTCTGTTTCAGTAATTTTCAGTTAAAGAA 180
 Qy 55 SerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu----- 69
 Db 181 ACCAACTTTCATAAAACATCTCATAGCTCTTAAACAAGATGTTGATTCGTGAAAT 240
 Qy 70 ---LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsn 88
 Db 241 GATACCTGGTAGTAATACAGAGGTATTAGTAAACCATCTAACTTTACAACTCCCTTACT 300
 Qy 89 ProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLys 108
 Db 301 GCAGAATTTCAAGACACCACTACTAATGCAATGCTCTATTGGTTATGCTTCTGCTGAA 360
 Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn---AsnGly 127
 Db 361 GGTCCAGAATTTGAATAGAAATTAATCATATGAAATAATTTGATGTAAATAATCCACAGGG 420
 Qy 128 TyrLysArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
 Db 421 TATACTACAGTAAGAAGTCTTATAGATACCTTGTCTTAGCAGCGTAATAAATTTCT 480
 Qy 147 SerAsnAsnProAlaAsnAsnLys-----TyrValThrLeuIleAsnAsn 161
 Db 481 CTATTCCAAACCAACAAACAAAGAGTAGTGGAATTTACCATGCTCGTAATGAACACAT 540
 Qy 162 GlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHis 181

Db 541 GGGTTATCTATCTTATCCAATATAGTAAATATTTGCTACGATTTTCTTAAATAATT 600
 Qy 182 AsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLys 201
 Db 601 CCTATATCACCTTATTTATGCGGAGGAATGGGTATAAATGCCATAGAATCTTTGACGCT 660
 Qy 202 TyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGln 221
 Db 661 TTACATGTGAATTTGCTTATCAAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 Qy 222 ValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysIle 241
 Db 721 ATCACTTATTTATGATGATATATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 780
 Qy 242 ProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHisVal 261
 Db 781 AAGTCCAA-----CATGTA 795
 Qy 262 His-----ThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
 Db 796 CATGAACCTTAAGATAATCCAAAGCTCACATCTGCAGTTGCTACACTTGATATAGCATAT 855
 Qy 274 TyrGlyGlySerIleGlyLysPheIle 283
 Db 856 TTTGGTAGTGAAGCTGGCATAAGAATTATA 885
 RESULT 11
 AAX34753
 ID AAX34753 standard; DNA; 888 BP.
 XX AAX34753;
 XX 05-JUL-1999 (first entry)
 XX DNA encoding OMP-LU protein.
 DE Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
 KW detection; dog; ss.
 XX Ehrlichia chaffeensis.
 OS WO9913720-A1.
 XX 25-MAR-1999.
 PD 18-SEP-1998; 98WO-US19600.
 PF 19-SEP-1997; 97US-0059353.
 PR (OHIS) UNIV OHIO STATE.
 PA Ohashi N, Rikihisa Y;
 PI WPI; 1999-254290/21.
 DR P-PSDB; AAY06953.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 13A; 55pp; English.
 PS The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the p30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 888 BP; 323 A; 127 C; 132 G; 306 T; 0 other;
 SQ Alignment Scores:
 Pred. No.: 3.59e-37 Length: 888

Score: 414.50 Matches: 96
Percent Similarity: 49.50% Conservative: 54
Best Local Similarity: 31.68% Mismatches: 124
Query Match: 27.45% Indels: 29
DB: 20 Gaps: 6

US-09-846-808-1 (1-284) x AAX34753 (1-888)

```
QY 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuLeuPheThrPro 20
   |||:|||||: ||| |||: |||: |||: |||
Db 1 ATGACAAGAAATTAATTTGTAATGTTATATACATTTTGTATTCTTTCCCA 60
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 21 HisIleSerLeuAlaSerValLeuAsnAspHisAsn-----SerMetTyr 35
   ||| ||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 61 CTTAAGTCATTTACACATATGCAATAATAACACATCACTCAAAAAGTTGGATTG 120
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 36 ValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGlu 55
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 121 ATAAGTGGTCAATATAGCCCAAGTATTCCTCATTTCAAGAATTTTTCAGTAGA 180
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThr 75
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 181 ---GACAAAGTAGATGTTGATGGTCTTACAACTGATGTATACATATACAGAACAT 237
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 76 GlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArg 95
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 238 ATATTACGAGATAATAACAAATTCACACACTCATATATTGCAAGTTCAAGAACATTTT 297
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 96 LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGlu 115
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 298 ATAAATTTACAGAGTGCATTTGGTTATTATCTGGCAAGGACCAAGGTTAGAAATAGAA 357
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 116 LeuSerTyrGluThrPheHisIleLysAsn---AsnGlyTyrLysArgIleAspCysGlu 134
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 358 AGCTCTTATGGGATTTGATGTTGTAAATATATAAAATATGACAGTACAAAGATGTTAAT 417
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 135 LysHisPheAlaLeuAlaLysGluIleSerGlyGlySer----- 147
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 418 AGATATTTTCTGTTAGTACGTGAAAAAATGGTTCAAAATTTCTCCAAAACCAACATGAA 477
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 148 -----AsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsn 161
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 478 ACTAGTCAACCCCTCGACAGTAATCTCTAAAAGTCTTTTATATCTTTAATGAAGATAAT 537
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 162 GlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHis 181
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 538 GGGGTATTTGTCATCAGTAATAATCAACGGTTGTATGATTTTCTTTTATAACACA 597
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 182 AsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLys 201
   ||| ||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 598 ACAATATCACCTTACGTTATGATAGGAGTTGGAGGAGATTTTATAGAGTTTGTGAAGTA 657
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 202 TyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGln 221
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 658 ATGCATATCAAGTTGCTGCCAAGTAAGTTGGTATTAGCTATCCAATATCTCCCTCT 717
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 222 ValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIle 241
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 718 ATTACTATTTTCTGCTGATGCATATATACAAAGTCAATAATAATTAATTAACACCTTA 777
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 242 ProValAsnTyrProCysAspTyrPro---SerProThrProProAsnSerLysProHis 260
   ||| ||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 778 CATGTTAAGTATTCATATGATGAATTAATAAACTACCTACC----- 816
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIle 280
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 817 ---ATTACCTCTGCAACAGCCAACTAAACATTTGAATATATTTTGGTGGTGAAGTTGGATG 873
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 281 LysPheIle 283
   :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 874 AGATTTATA 882
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
```

RESULT 12

AAC68705
ID AAC68705 standard; DNA; 843 BP.
XX
AC AAC68705;
XX
XX 02-MAR-2001 (first entry)
DT
XX
DE Ehrlichia chaffeensis VSA4 gene partial coding sequence.
XX
XX Ehrlichia chaffeensis; VSA4; variable surface antigen 4; MAP1;
KW major antigenic protein 1; antirickettsial; vaccine; gene therapy;
KW Rickettsia; Cowdria; Anaplasma; map2; lhworf3; 4hworf1; 18hworf1;
KW 3gdorf3; ds.
XX
OS Ehrlichia chaffeensis.
XX
PN WO200065063-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10886.
XX
PR 22-APR-1999; 99US-0130725.
XX
XX (UYFL) UNIV FLORIDA.
PA
XX Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmire WW, Alleman AR;
PI
DR WPI; 2000-679675/66.
DR P-PSDB; AAB36188.
XX
PT New polynucleotides useful as DNA vaccines for conferring immunity to
PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
PT
XX
XX Claim 4; Page 41; 63pp; English.
XX
CC The present sequence shows a high degree of similarity to the major
CC antigenic protein 1 (MAP1) gene of Ehrlichia sp. The MAP1 gene may be
CC used in a vaccine to protect animals or humans against rickettsial
CC diseases caused by a organisms of Rickettsia sp., Ehrlichia sp.,
CC Anaplasma sp. or Cowdria sp. The vaccine elicits an immune response
CC protective against the rickettsial pathogen. The nucleic acid vaccines
CC can be driven by the human cytomegalovirus (HCMV) enhancer-promoter
CC Cowdria ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1
CC and 3gdorf3 may be used in therapeutic and diagnostic applications. The
CC polypeptides are useful for detecting antibodies associated with
CC infection by a rickettsial pathogen whilst the polynucleotides may be
CC used to detect the presence of rickettsial nucleic acids.
XX
SQ Sequence 843 BP; 282 A; 127 C; 159 G; 275 T; 0 other;

Alignment Scores:
Pred. No.: 7,26e-37 Length: 843
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 21 Gaps: 11

US-09-846-808-1 (1-284) x AAC68705 (1-843)

QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuLeuPheThrProHisIleSer 23
 :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 12 AAAAATTTTTTATAACAACACTACATAGTATCGCTAATGCTTCTTACCTGGAAATCA 71
 |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
 :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 72 TTTTCTGATCGATACAGACAGCAATGTTGGTGGTAAATTTCTATATACAGTGGGAATAT 131
 |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaAlaAsnThrVal 60

```
Db 132 GTACCAAGTGTTCACATTTTGGCGTATTCTCTGCTAAACAGGAAGAAATACACAATC 191
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 192 GGAGTATTGGATTAAAGCAGATGGATGGCAGCACAATATCTAAATAATCTCCAGAA 251
Qy 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 252 AATACATTAAAGTTCACAAATATTATTAATATGAAATAATCCATTTCTAGGTTT 311
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 312 GCAGGAGCTGTGTGTTATTAATGAAT---GGTCCAGAAATAGAGTTAGAAATGCTCTAT 368
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 369 GAACATTGTGATGGAACACAGGGTAAATACATATAGAAC---GATGCTCAAAATAT 425
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrVal 156
Db 426 TATGCTTTAAACCCATAAC---AGTGGGGAAAGCTAAGCAATGCAGGTGATAAGTTGTT 482
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 483 TTTCTAAAAAATGAAGACTACTTGTATATATCACTTATGTGTAATGCATGCTATGATGTA 542
Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 543 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTGCAGGTGTGGTACTGAT 596
Qy 195 ThrIleAspPheLeuSerLysTyrThrThrylPheSerTyrGlnGlyLysLeuGlyAla 214
Db 597 TTAATATCCATGTTTGAAGCTATAAACCTTAAATTTCTTATCAAGGAAAGTTAGGTTG 656
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 657 AGTTACTCCATAAGCCAGAGCTCTGTTTGTGTCGACATTTTTCATAAGGTGATA 716
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 717 GGGATGAATTCACA-----GATATCTCTGCTATGATACCC 752
Qy 255 ProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
Db 753 AGTACCTCAACTCTCACAGGTAATCACTTTACTATATAGTAACTAAGTATGCGCACTTT 812
Qy 275 GlyGlySerIleGlyIleLysPhe 282
Db 813 GGAGTGGAACTTGGAGGAAGGTTT 836

RESULT 13
AAV07179
ID AAV07179 standard; DNA; 4683 BP.
AC XX
AC XX
AC XX
AC XX
DT 14-SEP-1998 (first entry)
XX
XX Ehrlichia chaffeensis VSA genomic locus.
DE MAP1 homologue; variable surface antigen; VSA1; VSA2; VSA3; VSA4;
KW VSA5; rickettsia; DNA vaccine; ss.
XX
XX Ehrlichia chaffeensis.
OS
XX Key Location/Qualifiers
FH CDS 3..134
FT CDS /*tag= a
FT /*note= "VSA1 partial gene"
FT terminator 203..212
FT /*tag= b
FT terminator 226..239
```

```
FT /*tag= c
FT -35_signal 349..354
FT /*tag= d
FT -10_signal 375..380
FT /*tag= e
FT RBS 430..434
FT /*tag= f
FT CDS 439..4299
FT /*tag= g
FT /*note= "VSA2"
FT terminator 1300..1309
FT /*tag= h
FT terminator 1324..1333
FT /*tag= i
FT misc_feature 1347..1361
FT /*tag= j
FT /*note= "G-rich region"
FT -35_signal 1473..1478
FT /*tag= k
FT -10_signal 1499..1504
FT /*tag= l
FT RBS 1554..1558
FT /*tag= m
FT CDS 1563..2399
FT /*tag= n
FT /*note= "VSA3"
FT terminator 2444..2456
FT /*tag= o
FT terminator 2455..2477
FT /*tag= p
FT misc_feature 2491..2505
FT /*tag= q
FT -35_signal 2618..2623
FT /*tag= r
FT -10_signal 2644..2649
FT /*tag= s
FT RBS 2699..2703
FT /*tag= t
FT CDS 2708..3550
FT /*tag= u
FT /*note= "VSA4"
FT terminator 3586..3596
FT /*tag= v
FT terminator 3610..3622
FT /*tag= w
FT misc_feature 3586..3596
FT /*tag= x
FT /*note= "G-rich region"
FT -35_signal 3767..3772
FT /*tag= y
FT -10_signal 3788..3793
FT /*tag= z
FT RBS 3847..3851
FT /*tag= aa
FT CDS 3856..4683
FT /*tag= ab
FT /*note= "truncated VSA5 gene"
XX
XX WO9816554-A1.
XX
XX 23-APR-1998.
XX
XX 17-OCT-1997; 97WO-US19044.
XX
XX 17-OCT-1996; 96US-0733230.
XX (UYFL ) UNIV FLORIDA.
XX
XX Barbet AF, Burrigde MJ, Ganta RR, Mahan SM, McGuire TC;
XX Nyika A, Rurangirwa FR;
XX WPI; 1998-251232/22.
DR
```

DR P-PSDB; AAW51091-05.
 XX Composition containing nucleic acid encoding rickettsial antigen -
 PT useful for, e.g. stimulating protective immune response in humans or
 PT animals
 XX
 XX
 PS Claim 4; Fig 2A-B; 39pp; English.
 XX
 CC This is the DNA sequence of a 4.6 kb genomic locus of Ehrlichia
 CC chaffeensis that was obtained using a PCR cloning strategy based on
 CC identifying genes homologous to the major antigenic protein MAP1
 CC (see AAW51088) of Cowdria ruminantium. It includes 5 very similar
 CC but non-identical open reading frames (ORFs), of which ORF1 is a
 CC partial gene and ORF5 is nearly complete but lacks 5-7 amino acid
 CC codons and a termination codon (see AAW51091-95). Due to their
 CC similarity to MAP1 surface antigen genes of C. ruminantium, the
 CC E. chaffeensis ORFs are designated variable surface antigen (VSA)
 CC genes 1-5. A claimed composition comprises a nucleic acid (see
 CC AAW07176-82) encoding a polypeptide (see AAW51088-99) that elicits a
 CC protective immune response against a rickettsial pathogen. The
 CC nucleic acid is used, in human or veterinary medicine, in vaccines
 CC to protect against Rickettsia, Ehrlichia, Anaplasma and Cowdria
 CC species. The nucleic acids are also useful as probes to identify
 CC related sequences, e.g. for identification of organisms and for
 CC diagnosing infection. Use of nucleic acid vaccines avoids the
 CC problem of protein purification associated with protein-based
 CC vaccines. The nucleic acid does not replicate in the host but
 CC remains episomal and capable of expressing polypeptide for at least
 CC 19 mth.
 XX
 SQ Sequence 4683 BP; 1590 A; 753 C; 818 G; 1522 T; 0 other;

Alignment Scores:

Pred. No.: 8,61e-36 Length: 4683
 Score: 411.50 Matches: 103
 Percent Similarity: 52.08% Conservative: 47
 Best Local Similarity: 35.76% Mismatches: 113
 Query Match: 27.25% Indels: 25
 DB: 19 Gaps: 11

US-09-846-808-1 (1-284) x AAW07179 (1-4683)

QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
 DB 2717 AAAAAATTTTATACAACTACATAGTAGTCGTAATGCTCTTCTACCTGGAATATCA 2776
 QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
 DB 2777 TTTTCTGATGCGAGTACAGAACGACAAATGTGGTGGTAATTTCTATATCATCGGGAAATAT 2836
 QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
 DB 2837 GTACCAAGTGTTCACATTTTGGCGTATTTCTGCTAACAGGAAAGAAATACAAATC 2896
 QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
 DB 2897 GGAGTATTGGATTAAAGCAAGATGGGATGGCAGCACAAATATCTAAATAATCTCCAGAA 2956
 QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 DB 2957 AATACATTTTAACTGTCCTCAATTTATTTTAAATATGAAATAATATCCATTTCTAGTGT 3016
 QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuLeuSerTyr 118
 DB 3017 GCAGGAGCTGTGGTTATTATGAAT---GGTCCAAGATAGAGTAGAATATGCTCTAT 3073
 QY 119 GluThrPheHisLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
 DB 3074 GAAACATTTGATGTGAAAAACAGGCTAATAACTATAAGAAC---GATGCTCAAAATAT 3130
 QY 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
 DB 3131 TATGCTTTTAAACCATAAC---AGTGGGGGAAAGCTAAGCAATGCGAGGTGATAATTTGTT 3187

QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 DB 3188 TTTCTAAAAAATGAAGCACTATTGATATATCATCTTATGTTGAATGCATGCTATGATGTA 3247
 QY 177 -----AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAsp 194
 DB 3248 ATAAGTGAAGGAATACCTTTCTCT-----CCTTACATATGTCAGGTGTGTGACTGAT 3301
 QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
 DB 3302 TTAATATCATGTTTGAAGCTATAAACCCTAAAATTTCTTATCAAGGAAAGTTAGGTTG 3361
 QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
 DB 3362 AGTACTCCATAAGCCAGCAAGCTCTGTTTGGTGGACATTTTTCATAAGGTGATA 3421
 QY 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
 DB 3422 GGAATGAATTCAGA-----GATATTCTGCTATGATACCC 3457
 QY 255 ProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
 DB 3458 AGTACCTCAACTCTCACAGGTAATCACTTTACTATATAGTAACACTAAGTGTATGCCACTTT 3517
 QY 275 GlyGlySerIleGlyIleLysPhe 282
 DB 3518 GGAGTGAACCTTGGAGGAAGGTTT 3541
 RESULT 14
 AAC68716
 ID AAC68716 standard; DNA; 4683 BP.
 XX
 AC AAC68716;
 XX
 DT 02-MAR-2001 (first entry)
 XX
 DE Ehrlichia chaffeensis 28 kDa gene locus.
 XX
 KW Ehrlichia chaffeensis; MAP1; major antigenic protein 1; antirickettsial;
 KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; ihworf3;
 KW 4hworf1; 18hworf1; 3gdorf3; ds.
 XX
 OS Ehrlichia chaffeensis.
 XX
 PN WO200065063-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 21-APR-2000; 2000WO-US10886.
 XX
 PR 22-APR-1999; 99US-0130725.
 XX
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Barbet AF, Bowie MV, Ganta RR, Burrledge MJ, Mahan SM, McGuire TC;
 PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmire WW, Alleman AR;
 XX
 DR WPI; 2000-679675/66.
 XX
 PT New polynucleotides useful as DNA vaccines for conferring immunity to
 PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
 PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
 PT -
 XX
 PS Example 2; Fig 2A-2B; 63pp; English.
 CC
 CC The present sequence is given in a specification relating to nucleic
 CC acid vaccines which may be used to protect animals or humans against
 CC rickettsial diseases caused by a organisms of Rickettsia sp.,
 CC Ehrlichia sp., Anaplasma sp. or Cowdria sp. The vaccine elicits an
 CC immune response protective against the rickettsial pathogen. The
 CC vaccines comprises the major antigenic protein 1 (MAP1) or major

CC antigenic protein 2 (MAP2) of rickettsial pathogens. The nucleic acid
 CC vaccines can be driven by the human cytomegalovirus (HCMV)
 CC enhancer-promoter. Cowdria ruminatum genes designated map 2, ihworf3,
 CC 4hwof1, 10hwof1 and 3dofr3 may be used in therapeutic and diagnostic
 CC applications. The polypeptides are useful for detecting antibodies
 CC associated with infection by a rickettsial pathogen whilst the
 CC polynucleotides may be used to detect the presence of rickettsial
 CC nucleic acids.

XX Sequence 4683 BP; 1576 A; 747 C; 831 G; 1526 T; 3 other;

Alignment Scores:

Pred. No.: 8,61e-36 Length: 4683
 Score: 411.50 Matches: 103
 Percent Similarity: 52.08% Conservative: 47
 Best Local Similarity: 35.76% Mismatches: 113
 Query Match: 27.25% Indels: 25
 DB: 21 Gaps: 11

US-09-846-808-1 (1-284) x AAC68716 (1-4683)

Qy 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
 Db 2717 AAAAAATTTTATACAACTACATTAGTATCGCTAATGTCCTTCTTACCTGGAATATCA 2776
 Qy 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
 Db 2777 TTTTCTGATGCAGTACAGAACGACAAATGTTGGTGGTAATTTCTATATCATGCGGAATAT 2836
 Qy 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
 Db 2837 GTACCAAGTCTTTCACATTTTGGCGTATTCTCTGCTAAACAGCAAGAAATAACAATC 2896
 Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
 Db 2897 GGAGTATTGGATTAAGCAAGATCGGATGGCAGCACAAATATCTAAAAATTCACAGAA 2956
 Qy 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 Db 2957 AATACATTTTAACGTTCCAAATATTATTCATTAAATAGAAATAATCCATTTCTAGGTTT 3016
 Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 3017 GCAGGAGCTGTTGGTTATTAAATGAAT---GGTCCAGAAATAGAGTAGAATGTCCTAT 3073
 Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
 Db 3074 GAACATTTTGATGTGAAACACAGGTAATAACTATAAGAAC---GATGCTCACAAATAT 3130
 Qy 137 PheAlaLeuAlaLysGluIleSerGlyLysSerAsnAsnProAlaAsnLysTyrVal 156
 Db 3131 TATGCTTTTAACCCATAAC---AGTGGGGGAAAGCTAAGCAATGCAGGTAGTAAAGTTGT 3187
 Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 Db 3188 TTTCTAAAAAATGAAGGACTACTTGATATATCACTTATCTTGAATGCATGCTATGATGA 3247
 Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValasp 194
 Db 3248 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTCAGGTGTGGTACTGAT 3301
 Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
 Db 3302 TTAATATCCATGTTTGAAGCTATAAACCCCTAAATTTCTTATCAAGAAAGTAGGTTTG 3361
 Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
 Db 3362 AGTTACTCCATAAGCCAGCAAGCTTCTGTTTGTGGTGGCAATTTTTCATAAGGTGATA 3421
 Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
 Db 3422 GGGAAATGAATTCAGA-----GATATTCCTGCTATGATACCC 3457

Qy 255 ProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
 Db 3458 AGTACCTCAACTCTCACAGGTAAATCATTACTATAGTAACTAAGTGTATGCCACTTT 3517
 Qy 275 GlyGlySerIleGlyIleLysPhe 282
 Db 3518 GGAGTGGAACTGGGAGGAAGGTTT 3541
 RESULT 15
 AAS07578
 ID AAS07578 standard; DNA; 4683 BP.
 XX AAS07578;
 XX 23-OCT-2001 (first entry)
 DE DNA encoding variable surface antigens 1-5 (VSA1-5) from E. chaffeensis.
 DE Major antigenic protein 1; MAP1; vaccine; immunogenic; rickettsial;
 KW infection; heartwater; diagnostic; variable surface antigen; VSA; ds.
 XX Ehrlichia chaffeensis.
 OS
 FH Key Location/Qualifiers
 CDS 3..134
 FT /*tag= a
 FT /product= "Variable surface antigen 1 (VSA1)"
 FT /partial
 FT /note= "No start codon"
 FT terminator 179..188
 FT /*tag= b
 FT /note= "Transcription terminator of VSA1 gene"
 FT terminator 202..212
 FT /*tag= c
 FT /note= "Transcription terminator of VSA1 gene"
 FT GC_signal 226..239
 FT /*tag= d
 FT -35_signal 349..354
 FT /*tag= e
 FT /note= "-35 region of VSA2 gene"
 FT -10_signal 375..380
 FT /*tag= f
 FT /note= "-10 region of VSA2 gene"
 FT RBS 430..434
 FT /*tag= g
 FT /note= "Ribosome binding site of VSA2 gene"
 FT CDS 439..1299
 FT /*tag= h
 FT /product= "Variable surface antigen 2 (VSA2)"
 FT terminator 1300..1309
 FT /*tag= i
 FT /note= "Transcription termination signal of VSA2"
 FT terminator 1324..1333
 FT /*tag= j
 FT /note= "Transcription termination signal of VSA2"
 FT GC_signal 1349..1361
 FT /*tag= k
 FT /note= "GC-rich sequence of VSA3 gene"
 FT -35_signal 1473..1478
 FT /*tag= l
 FT -10_signal 1499..1504
 FT /*tag= m
 FT /note= "-10 region of VSA3 gene"
 FT RBS 1554..1558
 FT /*tag= n
 FT /note= "Ribosome binding site of VSA3 gene"
 FT CDS 1563..2399
 FT /*tag= o
 FT /product= "Variable surface antigen 3 (VSA3)"
 FT terminator 2443..2455
 FT /*tag= p

```
FT /note= "Transcription termination signal of VSA3"
FT 2465..2477
FT /*tag= q
FT /note= "Transcription termination signal of VSA3"
FT 2491..2505
FT /*tag= r
FT /note= "GC-rich sequence of VSA4 gene"
FT 2618..2623
FT /*tag= s
FT /note= "-35 region of VSA4 gene"
FT 2644..2649
FT /*tag= t
FT /note= "-10 region of VSA4 gene"
FT 2699..2703
FT /*tag= u
FT /note= "Ribosome binding site of VSA4 gene"
FT 2708..3550
FT /*tag= v
FT /product= "Variable surface antigen 4 (VSA4)"
FT 3586..3596
FT /*tag= w
FT /note= "Transcription termination signal of VSA4"
FT 3610..3622
FT /*tag= x
FT /note= "Transcription termination signal of VSA4"
FT 2491..2505
FT /*tag= y
FT /note= "GC-rich sequence of VSA5 gene"
FT 3767..3772
FT /*tag= z
FT /note= "-35 region of VSA5 gene"
FT 3788..3793
FT /*tag= aa
FT /note= "-10 region of VSA5 gene"
FT 3847..3851
FT /*tag= ab
FT /note= "Ribosome binding site of VSA5 gene"
FT 3856..4683
FT /*tag= ac
FT /product= "Variable surface antigen 5 (VSA5)"
FT /partial
FT /note= "No stop codon"
XX
XX USG251872-B1.
XX
XX 26-JUN-2001.
XX
XX 17-OCT-1997; 97US-0953326.
XX
XX 17-OCT-1996; 96US-0733230.
XX
XX (UYFL ) UNIV FLORIDA.
XX
XX Barbet AF, Ganta RR, McGuire TC, Burrige MJ, Nyika A;
XX Rurangirwa FR, Mahan SM, Bowie MV, Alleman AR;
XX
XX WPI; 2001-424487/45.
XX
XX P-PSDB; AAU04195, AAU04196, AAU04197, AAU04198, AAU04199.
XX
XX New MAP2 genes and polypeptides useful as vaccines for conferring
XX immunity to human and animal rickettsial diseases, e.g. heartwater, or
XX as molecular markers in nucleic acid analysis procedures -
XX
XX Example 3; Fig 2A-2B; 30pp; English.
XX
XX The sequence represents the coding sequence of variable surface antigen
XX (VSA) gene locus encoding VSA1-5 proteins of Ehrlichia chaffeensis, which
XX have similarity to major antigen proteins (MAP). The MAP polynucleotides
XX and polypeptides are useful as vaccines for conferring immunity to
XX rickettsia infection, including Cowdria ruminantium causing heartwater.
XX The MAP polynucleotides may be used as molecular markers in nucleic acid
XX analysis procedures, and to produce the MAP polypeptides, which may
XX be used to raise antibodies that are reactive with the polypeptides.
XX
```

```
CC The nucleic acids may further be used as probes to identify
CC complementary sequences within other nucleic acid molecules or genomes,
CC where such probes can be applied to identify or distinguish infectious
CC strains of organisms in diagnostic procedures or in rickettsial
CC research where identification of particular organisms or strains is
CC needed.
XX
XX SQ Sequence 4683 BP; 1578 A; 739 C; 833 G; 1533 T; 0 other;
```

Alignment Scores:

```
Pred. NO.: 8.61e-36 Length: 4683
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 22 Gaps: 11
```

US-09-846-808-1 (1-284) x AAS07578 (1-4683)

```
Qy 7 ArgLysPheValLeu-----TrpValMetLeuLeuLeuPheThrProHisIleSer 23
Db 2717 AAAAAATTTTATTAACAACACTACATTAGTAGCGCTAATGTCCTCTTACCTGGAATATCA 2776
Qy 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 2777 TTTTCTGATGCAGTACAGACAGCAATGTTGGTGCTAATTTCTATATCAGTGGGAATAT 2836
Qy 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaAlaAsnThrVal 60
Db 2837 GTACCAAGTGTTCACATTTTGGCGTATTCTCTGCTAAACAGGAAGAAATACACAATC 2896
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 2897 GGAGTATTGGATTAAACAAGATGGGATGGCACACACAAATATCTAAAAATCTCCAGNA 2956
Qy 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 2957 AATACATTAAACGTTCCAAATATTATTAATATGAAATAATAATCCATTTCTAGGTTT 3016
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 3017 GCAGGAGCTGTGGTTATTATTAAATGAAT---GGTCCAAGAANTAGAGTTAGAAATGTCCTAT 3073
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 3074 GAAACATTTGATGTGAAAACACAGGGTAAATAACTATAAGAAC---GATGCTCACAAATAT 3130
Qy 137 PheAlaLeuAlaLysGluLysSerGlySerAsnAsnProAlaAsnLysTyrVal 156
Db 3131 TATGCTTTAAACCCATAAC---AGTGGGGGAAAGCTTAAGCAATGCAGGTGATAAGTTGTT 3187
Qy 157 ThrLeuLeuAsnAspGlyIleSerLeuThrSerAlaLeuLeuAsnValCystTyrAspVal 176
Db 3188 TTTCTAAAAAATGAAGGACTACTTGATATATCACTATTGTTGAATGCATGCTATGATGA 3247
Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 3248 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTGCAGGTGTTGGTACTGAT 3301
Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 3302 TTAATATCCATGTTTGAAGCTATAACCCCTAAATAATTTCTTCAAGGAAGTAGGTGTTG 3361
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 3362 AGTTACTCCATAAGCCCAAGAGCTTCTGTTTGTGTGGACATTTTCATCAAGTGATA 3421
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 3422 GGGATGAATTCAGA-----GATATTCCTGCTATGATACCC 3457
Qy 255 ProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
```

Db	3458	AGTACCTCAACTCTCACAGTAAATCACATTACTATAGTACACAACTAAGTGATGCCACTTT	3511
Qy	275	Glycylserilecylileuysphe 282	
		:: ::	
Db	3518	GGAGTGAACCTGGAGGAAGGTTT	3541
RESULT 16			
	AAX34748		
ID	AAX34748 standard; DNA; 843 BP.		
XX	AAX34748;		
AC	AC		
XX	05-JUL-1999 (first entry)		
DT	DNA encoding OMP-1F protein.		
DE	DE		
XX	Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;		
KW	detection; dog; ss.		
KW	Ehrlichia chaffeensis.		
OS	WO9913720-A1.		
XX	25-MAR-1999.		
PN	18-SEP-1998; 98WO-US19600.		
XX	19-SEP-1997; 97US-0059353.		
XX	(OHIS) UNIV OHIO STATE.		
PA	Ohashi N, Rikihisa Y;		
XX	WPI: 1999-254290/21.		
DR	P-PSDB; AAY06948.		
DR	Novel outer membrane proteins from Ehrlichia chaffeensis and		
XX	Ehrlichia canis		
PT	Disclosure; Fig 8A; 55pp; English.		
XX	XX		
PS	XX		
CC	XX		
CC	The invention provides isolated outer membrane proteins (OMP) from		
CC	Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form pa		
CC	of the OMP family and consist of proteins OMP-1, -1(B to Z) shown		
CC	in AAY06943-958. The E. canis proteins form part of the P30 family and		
CC	consist of proteins shown in AAY06959-970. The proteins and genes are		
CC	used to detect E. chaffeensis in patients and E. canis in dogs.		
XX	XX		
SQ	Sequence 843 BP; 283 A; 128 C; 159 G; 273 T; 0 other;		
Alignment Scores:			
Pred. No.:	1-59e-36	Length:	843
Score:	408.50	Matches:	103
Percent Similarity:	51.74%	Conservative:	46
Best Local Similarity:	35.76%	Mismatches:	114
Query Match:	27.05%	Indels:	25
DB:	20	Gaps:	11
US-09-846-808-1 (1-284) x AAX34748 (1-843)			
Qy	7	ArglyspheValLeu-----TrpValMetLeulleLeuPheThrProHisIleSer	23
		::: :::	
Db	10	AAAAAATTTTATACAACTACATTAGTAGTCGCTAATGTCCTTCTTACCTGGAATATCA	69
Qy	24	LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr	40
		::: ::: :::	
Db	70	TTTTCTCATGCAGTACAGAACACAAATGTTGGTGTAATTTCTATATCATCGTGGAAATAT	129
Qy	41	LysProAlaArgGlnHisLeuSerLysLeulleLeuLeuAlaAlaAsnThrVal	60
		: :::	
Db	130	GTACCAAGGTTTCACATTTTGGCGTATTCTCTGCTAAACAGGAAGNAATAACAACACC	189
Qy	61	GlulValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn	80

Db	190	GGAGTATTGGATTAAAGCAAGATTTGGGATGGCAGCACATATCTAAAAATTTCTCCAGAA	249
Qy	81	ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe	98
Db	250	AATACATTAAACGTGCCAAATTTATTCATTTAAATATGAAATAATCAATTTCTAGTGT	309
Qy	99	SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118
Db	310	GCAGGAGCTGTGGTTATTTAAATGAAT---GGTCCAAGATAGAGTTAGAAATGTCCTAT	366
Qy	119	GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis	136
Db	367	GAACATTTGATGTGAAAAACCCAGGTAATACTATAAGAAC---GATGCTCACAATAT	423
Qy	137	PheAlaLeuAlaLysGluIleSerSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal	156
Db	424	TATGCTTTAACCCATAAC---AGTGGGGAAAGCTAAGCAATCCAGGTGATAGATTGTT	480
Qy	157	ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal	176
Db	481	TTTCTAAAAAATGAAGGACTACTTGATATATACATTATGTTGAATGCATGCTATGATGA	540
Qy	177	-----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp	194
Db	541	ATAAGTGAAGGAATACCTTTCTCT-----CCTTACATATGTCGAGGTGTGTACTGAT	594
Qy	195	ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla	214
Db	595	TTAATATCCATGTTGAAGCTATAAACCCCTAAAAATTTCTTATCAAGAAAGTAGGTTTG	654
Qy	215	SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe	234
Db	655	AGTTACTCCATAAGCCCGAGAGCTTCTGTTTGTGGTGACATTTTCATAAGGTGATA	714
Qy	235	GlyTyrLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro	254
Db	715	GGGAATGAATTCAGA-----GATATTCTGCTATGATACCC	750
Qy	255	ProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyr	274
Db	751	AGTACCTCAACTCCACAGGTATACACTTTACTATAGTACACTAGTGTATGCCACTTT	810
Qy	275	GlyGlySerIleGlyIleLysPhe	282
Db	811	GGAGTGAACCTTGAGGAAGGTTT	834
RESULT	17		
AAAD01294			
ID	AAAD01294	standard; DNA; 840 BP.	
AC	AAAD01294;		
DT	12-OCT-2000	(first entry)	
DE	Ehrlichia canis immunoreactive protein Eca28SA3	DNA.	
KW	Homologous 28-kDa protein gene; Eca28SA3; immunoreactive; vaccine;		
KW	p28 gene; polymorphic multiple gene family; immunoprotective antigen;		
KW	antibacterial; canine ehrlichiosis; canine tropical pancytopenia;		
KW	tick-borne rickettsial disease; serodiagnostics; ds.		
OS	Ehrlichia canis.		
XX	Key	Location/Qualifiers	
FT	CDS	1..840	
FT		/*tag= a	
FT		/product= "Eca28SA3 protein (30-kDa)"	
FT		/note= "Does not include stop codon"	
FT		/partial	
FT	sig_peptide	1..69	
FT		/*tag= b	
FT	mat_peptide	70..840	

FT /*tag= c
 FT /product= "Mature Eca28SA3 protein (28-kDa)"
 PN
 XX WO200032745-A2.
 XX

PD 08-JUN-2000.
 XX

XX 24-NOV-1999; 99WO-US28075.
 XX

XX 30-NOV-1998; 98US-0201458.
 PR

XX 03-MAR-1999; 99US-0261358.
 PR

XX (RERE-) RES DEV FOUND.
 XX

XX Walker DH, Yu X, McBride JW;
 PI

XX WPI; 2000-412298/35.
 DR

XX P-PSDB; AAY1479.
 DR

XX Ehrlichia canis antigens useful for vaccinating against canine
 PT ehrlichiosis in dogs -
 PT

XX Claim 5; Page 67-68; 86pp; English.
 PS

XX The patent relates to homologous 28-kilodalton (kDa) protein genes of
 CC Ehrlichia canis, designated Eca28SA1, Eca28SA2, Eca28SA3, Eca28-1 and
 CC Eca28-2. These genes are members of a polymorphic multiple gene family
 CC and contained in a single locus of 5.592 kb. The 28-kDa proteins are
 CC immunoreactive with anti-E. canis serum hence are important
 CC immunoprotective antigens. The protein is useful for vaccinating
 CC against E. canis infections such as canine ehrlichiosis in dogs.
 CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
 CC tick-borne rickettsial disease of dogs. Eca28-1 is conserved amongst
 CC different strains of E. canis and hence useful for serodiagnosis of
 CC canine ehrlichiosis. The present sequence is a DNA encoding E. canis
 CC Eca28SA3 30-kDa protein which is post-translationally modified to a
 CC mature 28-kDa protein by cleavage of N-terminal signal sequence.
 XX

SQ Sequence 840 BP; 282 A; 137 C; 150 G; 271 T; 0 other;

Alignment Scores:

Pred. No.: 2,53e-34 Length: 840
 Score: 389,00 Matches: 98
 Percent Similarity: 51.09% Conservative: 43
 Best Local Similarity: 35.51% Mismatches: 115
 Query Match: 25.76% Indels: 20
 DB: 21 Gaps: 10

US-09-846-808-1 (1-284) x AAD01294 (1-840)

QY 15 LeuileuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn----- 32
 DB 43 TTAATGACTATGCTCCAGCATATCTTTCTGATCTATACAGAGCATACACCTGGT 102
 QY 33 SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuIle 52
 DB 103 AGCTTCTACATCAGTGGAAATATATGACCAAGTCTTTCACATTTTGGTGTTCAGCT 162
 QY 53 LysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuAsnAsp 72
 DB 163 AAAGAAGAAAGAAACTCAACTGTGGAGTGTGGATTAACATGATGGAGGT 222
 QY 73 LeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLys---TyrAsnProTyrTyr 91
 DB 223 ACAATATCTAATCTTCTCCAGAAATATATTCACATTCATAAATATTCGTTAAATAC 282
 QY 92 GluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnPhe 110
 DB 283 GAAACAACCACTCTTCTAGGTTTGCAGAGCTATTGTTATTCAATGGGT---GGCCCA 339
 QY 111 ArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----AsnAsnGlyTyr 128
 DB 340 AGAATAGAACTTGAAGTCTGTACGAGACATTCGATGTGAAATATCAGAAATAATAT 399

QY 129 LysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsn 148
 DB 400 AAGAAC---GGCGCACACAGATACCTGCTTTATCTCATCATAGTTACAGCAACAGCATG 456
 QY 149 AsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAla 168
 DB 457 TCCTCCGCAAGTAACAATTTGTTTCTTAAAAAATGAAGGTTAATTGACTTATCATTT 516
 QY 169 LeuIleAsnValCysTyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyr 186
 DB 517 ATGATAAATGCATGCTATGACATAATAATTGAAGGAATGCCTTTTCA-----CCTTAT 570
 QY 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPhe 206
 DB 571 ATTGTGCAGGTGTGGTACTGATGTTGTTCCATGTTTGAAGCTATAAACTCTAAATTT 630
 QY 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
 DB 631 TCTTACCAAGGAAACTAGGATTAGGTTATAGTATAAGTTTCAGAGGCTCTGTTTTATTC 690
 QY 227 GluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysIleProValAsnTyrPro 246
 DB 691 GGTGGACACTTTCACAGAGTCATAGGTAATGAATTTAGAGACATCCCTGCTATGTTCT 750
 QY 247 CysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThrAlaLeu 266
 DB 751 AGTGA-----TCAATCTTCCAGAAACCAA-----TTTCAATA 786
 QY 267 AlaMetLeuSerIleGlyTyrTyrGlySerIleGlyLysPhe 282
 DB 787 GTAACACTAAATGTGTCTACCTTTGGCATAGAACTTGGAGGAAGATTT 834
 RESULT 18
 AAD01295
 ID AAD01295 standard; DNA; 2037 BP.
 XX
 AC AAD01295;
 XX
 DT 12-OCT-2000 (first entry)
 XX
 DE Ehrlichia canis immunoreactive protein genes Eca28SA2 and Eca28SA3.
 KW Homologous mature 28-kDa protein gene; Eca28SA2; Eca28SA3; vaccine;
 KW p28 gene; polymorphic multiple gene family; immunoprotective antigen;
 KW antibacterial; canine ehrlichiosis; canine tropical pancytopenia;
 KW tick-borne rickettsial disease; serodiagnosis; immunoreactive; ds.
 XX
 OS Ehrlichia canis.
 XX
 PH Key Location/Qualifiers
 FT CDS 1..852
 FT /*tag= a
 FT /product= "Eca28SA2 protein (30-kDa)"
 FT misc_feature 853..1194
 FT /*tag= b
 FT /note= "Intergenic non-coding region NC2"
 FT CDS 1195..2037
 FT /*tag= c
 FT /product= "Eca28SA3 protein (30-kDa)"
 FT sig_peptide 1195..1263
 FT /*tag= d
 FT mat_peptide 1264..2034
 FT /*tag= e
 FT /product= "Mature Eca28SA3 protein (28-kDa)"
 XX
 XX WO200032745-A2.
 XX
 XX 08-JUN-2000.
 PD
 XX 24-NOV-1999; 99WO-US28075.
 PF
 XX 30-NOV-1998; 98US-0201458.
 PR
 XX 03-MAR-1999; 99US-0261358.
 PR

PR 03-MAR-1999; 99US-0261358.
 PA (RERE-) RES DEV FOUND.
 XX Walker DH, Yu X, McBride JW;
 XX WPI; 2000-412298/35.
 DR P-PSDB; AAY71478, AAY71479.
 XX Ehrlichia canis antigens useful for vaccinating against canine
 PT ehrlichiosis in dogs -
 XX
 PS Example 15; Fig 7; 86pp; English.
 XX
 CC The patent relates to homologous 28-kiloDalton (kDa) protein genes of
 CC Ehrlichia canis, designated Eca28SA1, Eca28SA2, Eca28SA3, Eca28-1 and
 CC Eca28-2. These genes are members of a polymorphic multiple gene family
 CC and contained in a single locus of 5.592 kb. The 28-kDa proteins are
 CC immunoreactive with anti-E. canis serum hence are important
 CC immunoprotective antigens. The protein is useful for vaccinating
 CC against E. canis infections such as canine ehrlichiosis in dogs.
 CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
 CC tick-borne rickettsial disease of dogs. Eca28-1 is conserved amongst
 CC different strains of E. canis and hence useful for serodiagnosis of
 CC canine ehrlichiosis. The present sequence is that of homologous
 CC genes encoding E. canis Eca28SA2 and Eca28SA3 30-kDa proteins which
 CC are post-translationally modified to corresponding
 CC mature 28-kDa proteins by cleavage of N-terminal signal sequence.
 XX
 SO Sequence 2037 BP; 687 A; 321 C; 342 G; 687 T; 0 other;
 Alignment Scores:
 Pred. No.: 9,06e-34 Length: 2037
 Score: 389.00 Matches: 98
 Percent Similarity: 51.09% Conservative: 43
 Best Local Similarity: 35.51% Mismatches: 115
 Query Match: 25.76% Indels: 20
 DB: 21 Gaps: 10
 US-09-846-808-1 (1-284) x AAD01295 (1-2037)
 Qy 15 LeuileLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn----- 32
 Db 1237 TTAATGTACTACTGCTCCAGCAATATCTTTTCGTGATACATATACAGACGAAACACTGGT 1296
 Qy 33 SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuile 52
 Db 1297 AGCTTCTACATCAGTGGAAATATGTACCAAGTGTTCACATTTTGGTGTCTCAGCT 1356
 Qy 53 LysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAsp 72
 Db 1357 AAAGAAGAAAGAACTCAACTGTTGGAGCTTTTGGATTAAACATGATTGGAATGGAGGT 1416
 Qy 73 LeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLys---TyrAsnProTyrTyr 91
 Db 1417 ACAATATCAATCTCTCTCCAGAAATATATTCACAGTTCAAAATATTATCGTTTAAATAC 1476
 Qy 92 GluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsnPhe 110
 Db 1477 GAAACAACCCATCTTAGGTGTTTCAGAGAGCTATTGGTTATTCATCGGT---GGCCCA 1533
 Qy 111 ArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----AsnAsnGlyTyr 128
 Db 1534 AGAATAGAAGTGAAGTCTGTACAGACATTCGATGCGAAGAAATCAGAACAAATTAATAT 1593
 Qy 129 LysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyClySerAsn 148
 Db 1594 AAGAAC---GGCGCACACAGATGCTGCTTATCTATCTATCATGATTCAGCAACAGCATG 1650
 Qy 149 AsnProAlaAsnAsnLysTyrValThrLeuLeuLeuAsnAsnGlyIleSerLeuThrSerAla 168
 Db 1651 TCCTCCGCAAGTAACAATTTGTTTCTTTAAATAAAGAGGTTTAATTCGACTTATCATTT 1710

Qy 169 LeuileAsnValCysTyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyr 186
 Db 1711 ATGATAAATGCATGCTATGACATAATAATTAAGGAATGCCCTTTTCA-----CCTTAT 1764
 Qy 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPhe 206
 Db 1765 ATTTGTGAGGTGTTGGTACTGATGTTTTCCTCATGTTTGAAGCTATAAATCCTAAAT 1824
 Qy 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
 Db 1825 TCTTACCAGGAAACTAGGATTAGTTATAGTATAAGTTTCAGAACCTCTGTTTATC 1884
 Qy 227 GluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrPro 246
 Db 1885 GGTGCACACTTTCACAGAGTCATAGTAAATTTAGAGACATCCTGCTATGGTTCCT 1944
 Qy 247 CysAspTyrProSerProThrProAsnSerLysProHisValHisThrThrAlaLeu 266
 Db 1945 AGTGA-----TCAAAATCTCCAGAAACCAA-----TTTGCAATA 1980
 Qy 267 AlaMetLeuSerIleGlyTyrTyrGlySerIleGlyIleLysPhe 282
 Db 1981 GTAACACTAAATGTGTGCTGCTGTCATGTCATGAGAACTGGAGGAAGATT 2028
 RESULT 19
 ID AAX34747 standard; DNA; 837 BP.
 XX AAX34747;
 AC AAX34747;
 DT 05-JUL-1999 (first entry)
 XX DNA encoding OMP-1E protein.
 DE Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX Ehrlichia chaffeensis.
 OS WO9913720-A1.
 PN 25-MAR-1999.
 PD 18-SEP-1998; 98WO-US19600.
 PF 19-SEP-1997; 97US-0059353.
 PR (OHIS) UNIV OHIO STATE.
 PA Ohashi N, Rikihisa Y;
 PI WPI; 1999-254290/21.
 DR P-PSDB; AAY06947.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 PS Disclosure; Fig 7A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 SQ Sequence 837 BP; 280 A; 130 C; 159 G; 268 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,71e-34 Length: 837
 Score: 387.50 Matches: 102
 Percent Similarity: 51.05% Conservative: 44
 Best Local Similarity: 35.66% Mismatches: 117

Query Match:	25..66%	Indels:	23
DB:	20	Gaps:	10
US-09-846-808-1 (1-284) x AAX34747 (1-837)			
Qy	7	ArgLysPheValLeu-----TrrValMeLeuIleLeuPheThrProHisIleSer	23
Db	10	AAAAAATTTTATAACAACGTGCAATAGTATCACTAATGTCCTTTCTACCTGGAATCA	69
Qy	24	LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr	40
Db	70	TTTTCTGATCGAGTGAAGGTGACAATATTAGTGGTAATTTCTATGTTAGTGGCAAGTAT	129
Qy	41	LysProAlaArgGlnHisLeuSerLysLeuIleLysGluSerAlaAlaAsnThrVal	60
Db	130	ATGCCAAGTGCTTCGCATTTTGGCAGCTTTTCTGCCAAGAAGAAAAAATCCTACTGTT	189
Qy	61	GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn	80
Db	190	GCATTTGATGGCTTAAACAAGATTGGGAAGGATTAGTCTCATCAAGTCACAATGATAAT	249
Qy	81	ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe	98
Db	250	---CATTTCAATAACAAGGTTATTCATTATAATATGAAATAAACCCATTTTATAGGTTT	306
Qy	99	SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118
Db	307	GCAGAGCTATTGGTTATTCAATGGGT---GGTCCAAGAGTAGAGTGTGAAGTGCTCAT	363
Qy	119	GluThrPheIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis	136
Db	364	GAACATTTGACGTTAAAAATCAGGTTAATAACTATAAAAAAT---GATGCTCACAGATAC	420
Qy	137	PheAlaLeuAlaLysGluIleSerGlySerAsnAsnPheAlaAsnLysTyrVal	156
Db	421	TGTGCTTTAGGTCACACAGACAACACGCGGAATA-----CCTAAAACTAGTAAATACGTA	474
Qy	157	ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal	176
Db	475	CTGTTAAAAACGGAAGATTGCTTGACATATCATTTATGCTAAATGCATGCTATGATATA	534
Qy	177	AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIle	196
Db	535	ATAAACGAGAGCATACCTTTGTCCTCTTACATATGTGCAGGTGTTGGTACTGATTTAATA	594
Qy	197	AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr	216
Db	595	TCCATGTTGAAGCTACAAAATCCATAAATTTCTTACCAAGGGAAGTTAGGTCTAAGTTAC	654
Qy	217	ThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPheGlyLys	236
Db	655	TCTATAACCCCAAGGCTCTGTATTTTATTTGGTGGACATTTTCTAAGGTGATAGGAAC	714
Qy	237	LysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProProAsn	256
Db	715	GAATTAGGGACATTCCTACTCTGAAGCATTTTGGTACGTCATCAGCTACTCCA-----	768
Qy	257	SerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly	276
Db	769	-----GATCTAGCAATAGTAACACTAAGTGTATGTCTATTTTGGGAATA	810
Qy	277	SerIleGlyIleLysPhe	282
Db	811	GAACCTGGAGAAGGTTT	828
RESULT 20			
AAC68704			
ID	AAC68704 standard; DNA; 837 BP.		
XX			
AC	AAC68704;		
XX			
DT	02-MAR-2001 (first entry)		
XX			

Qy 81 ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 Db 252 ---CATTTCAATCAACAGGGTTATTCAATTAATGAATAAATACCCATTATTAGGGTTT 308
 Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 309 GCAGAGGCTATTGGTTATTCATCAATGGGT---GGTCCAAGAGTAGAGTTTGAAGTCTCCTAT 365
 Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAsnAsnLysHis 136
 Db 366 GAAACATTGACGTTAAATACAGGGTAATACTATAAATAAT---GATGCTCAGATAC 422
 Qy 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
 Db 423 TGTGCTTTAGTCAACAGACAGCGGAATA-----CCTAAAACTAGTAAATACGTA 476
 Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 Db 477 CTGTTAAACCGAGGATTCCTGCATATCATTTATGCTAAATGATGCTATGATATA 536
 Qy 177 AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
 Db 537 ATAAACGAGACATACCTTTGCTCTTACATATGTCAGGTGTTGGTACTGATTTAATA 596
 Qy 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
 Db 597 TCCATGTTTGAAGCTACAAATCCTAAATTTCTTACCAAGGGAAGTTAGGTCTAAGTTAC 656
 Qy 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
 Db 657 TCTATAAACCCAGAGCTCTGTATTTATGTTGGGACATTTTCATAAGGTGATAGGAAC 716
 Qy 237 LysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProProAsn 256
 Db 717 GAATTTAGGACATTCCTACTCTGAAAGCAATTTGTTAGTCTACAGTACTCTCA- 770
 Qy 257 SerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
 Db 771 -----GATCTAGCAATAGTAACACTAAGTGTATGTCATTTTGGGAATA 812
 Qy 277 SerIleGlyIleLysPhe 282
 Db 813 CAACTTGGAGGAAGGTTT 830
 RESULT 21
 AAX34762
 ID AAX34762 standard; DNA; 843 bp.
 XX AC AAX34762;
 XX DT 05-JUL-1999 (first entry)
 XX DE DNA encoding P30-2 protein.
 XX KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 XX OS detection; dog; ss.
 XX PN Ehrlichia canis.
 XX PN WO9913720-A1.
 XX PD 25-MAR-1999.
 XX PF 18-SEP-1998; 98WO-US19600.
 XX PR 19-SEP-1997; 97US-0059353.
 XX PA (OHIS) UNIV OHIO STATE.
 XX PI Ohashi N, Rikihisa Y;
 XX WP: 1999-254290/21.
 DR P-PSDB; AAY06962.

XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 22A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 843 BP; 283 A; 136 C; 150 G; 274 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 5,54e-34 Length: 843
 Score: 286.00 Matches: 97
 Percent Similarity: 51.09% Conservative: 44
 Best Local Similarity: 35.14% Mismatches: 115
 Query Match: 25.56% Indels: 20
 DB: 20 Gaps: 10
 US-09-846-808-1 (1-284) x AAX34762 (1-843)
 Qy 15 LeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn----- 32
 Db 43 TTAATGTACTATGCTCCAAGCATATCTTTTCTGTACTATACAGAGCAGATAACTGGT 102
 Qy 33 SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIle 52
 Db 103 AGCTTCTACATCAGTGGAAATATATGACCAAGTGTTCACATTTTGGTGTCTCAGCT 162
 Qy 53 LysGluSerAlaAlaAsnThrValGlnValPheGlyLeuLysLysAspLeuLeuAsnAsp 72
 Db 163 AAAGAAGAAAGAACTCAACTGTTGGAGTTTGGATTAAACACATGATGAATGGAGCT 222
 Qy 73 LeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLys---TyrAsnProTyrTyr 91
 Db 223 ACAATATCTACTCTCTCCAGAAATATATTCACAGTCAAAATATTCGTTAAATAC 282
 Qy 92 GluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsnPhe 110
 Db 283 GAAACCAACCCATCTTTAGGTTTGGAGGACTATTTGGTATTTCATCAATGGGT---GGCCCA 339
 Qy 111 ArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----AsnAsnGlyTyr 128
 Db 340 AGAATAGAACTTGAAGTCTGTACGAGACATTCGATGTGAAAAATCAGAACATAATAT 399
 Qy 129 LysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlySerAsn 148
 Db 400 AAGAAC---GGCGCACACAGATCTGTCTTATCTCATCATAGTTCAGCAACAACATG 456
 Qy 149 AsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAla 168
 Db 457 TCCCTCCCAAGTAAACAAATTTGTTTCTTAAAAATGAAGGTTAAATGACTTATCATTT 516
 Qy 169 LeuIleAsnValCysTyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyr 186
 Db 517 ATGATAATGCATCTATGACATAATAATGAAGAAATGCCCTTTTCA-----CCITAT 570
 Qy 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLysPhe 206
 Db 571 ATTTGTGCGGTGTTGGTACTGATGTTGTTCCATGTTTGAAGCTATATAAATCCTAAAT 630
 Qy 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
 Db 631 TCTTACCAAGGAAACTAGGATTAGTTATAGTATAAGTTTCAAGGCTCTGTTTATC 690
 Qy 227 GluGlyTyrHisGlyLeuPheGlyLysLysPheGlyLeuLysIleProValAsnTyrPro 246
 Db 691 GGTGGACACTTTTCACAGATCATAGGTAATGAATTTAGAGACATCCCTGCTGCTCCT 750

QY 247 CysAspTyrProSerProThrProAsnSerLysProHisValHisThrThrAlaLeu 266
 Db 751 AGNGGA-----TCAATCTCCAGAAACCAA-----TTTGCATA 786
 QY 267 AlaMetLeuSerIleGlyTyrGlyGlySerIleGlyIleLysPhe 282
 Db 787 GTAAACACATAATGTGTCACTTGTGTTAGAACTTGGAGGAAGATT 834

RESULT 22

AA34759
 ID AAX34759 standard; DNA; 867 BP.
 AC AAX34759;
 XX
 XX 05-JUL-1999 (first entry)
 XX
 XX DNA encoding P30 protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX
 XX Ehrlichia canis.
 XX
 XX WO9913720-Al.
 XX
 XX 25-MAR-1999.
 XX
 XX 18-SEP-1998; 98WO-US19600.
 XX
 XX 19-SEP-1997; 97US-0059353.
 XX
 XX (OHIS) UNIV OHIO STATE.
 XX
 XX Ohashi N, Rikihisa Y;
 XX
 XX WPI; 1999-254290/21.
 DR P-PSDB; AAY06959.
 XX
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 PT
 XX
 PS Disclosure; Fig 19A; 55pp; English.

CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX

SQ Sequence 867 BP; 305 A; 150 C; 157 G; 255 T; 0 other;

Alignment Scores:
 Pred. No.: 5,77e-34 Length: 867
 Score: 386.00 Matches: 99
 Percent Similarity: 50.70% Conservative: 46
 Best Local Similarity: 34.62% Mismatches: 113
 Query Match: 25,56% Indels: 28
 DB: 20 Gaps: 11

US-09-846-808-1 (1-284) x AAX34759 (1-867)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
 Db 37 ATATCACTAATGTCTTCTTACCTACCGGTATCTTTTCTGGAATCAATACATGAAGATAAT 96
 QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 Db 97 ATAAATGGTAACCTTTTACATTAGTGCAAAGTATATGCCAAGTGCCTCACACTTTGGCGTA 156
 QY 50 LeuLeuIleLysGlnSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp--- 68
 Db 157 TTTTCAGTTAAAGAGAGAGAAAAACACAACACTGGAGTTTCGGATTAAACAAAGATTGG 216

QY 69 -----LeuLeuAsnAspLeuLeuThrGly-----IleLysAspAsnThrAsnPhe 83
 Db 217 GACGGAGCAACAATAAAGGATGCAAGCAGCCACACACATAGACCCACCAATATTC 276
 QY 84 AsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIle 101
 Db 277 TCCATTTCAAATATTTCATTTAAATATGAACAATCCATTTTAGGGTTTGCAGGAGCT 336
 QY 102 PheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPhe 121
 Db 337 ATTGGCTACTCAATGGGT---GGTCCAAGGTAGAGTTTGAAGTGTCTTACGAATAATT 393
 QY 122 HisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeu 139
 Db 394 GATGTAATAAACCAAGGTAACAGTTACAAGAAC---GATGCTCACAATATTCGCGCTTTA 450
 QY 140 AlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIle 159
 Db 451 TCAAGACACACCGGAGGTATGCCACAAGCCGGTCATCAAAATAAATTTGTCTTCTTAAAA 510
 QY 160 AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeu 179
 Db 511 AATGAAGGATTACTTGACATATCACTTATGATAAACGCATGTTATGATATAACAATCGAC 570
 QY 180 LysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeu 199
 Db 571 AGCATGCCATTTCTCCATATATATGCGAGGTATGCGTAGTGACITAGTTTCGATGTTT 630
 QY 200 SerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSer 219
 Db 631 GAAACTACAAATCCTAAATTTCTTATCAAGGAAATTTAGGTGTGAAGTTACTCCATAAGC 690
 QY 220 ProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlu 239
 Db 691 CCAGAAGCATCTCTTTTGTGAGGACACTTTTCACAGAGTTTATAGGTAATGAATTTAAA 750
 QY 240 LysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSer----- 257
 Db 751 -----GACATTCCTGCAATAACTCTCTGCTGGAGACACAGAA 786
 QY 258 ---LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
 Db 787 ATTAAGGACACACAGTTTACAACA-----GTAACATTAAACATATGCCACTTCGGACTA 840
 QY 277 SerIleGlyIleLysPhe 282
 Db 841 GAGCTTGGAGGCAGGTTT 858
 RESULT 23
 ID AAV07176
 XX AAV07176 standard; DNA; 864 BP.
 XX
 XX AAV07176;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 XX Cowdria ruminantium MAP1 gene coding sequence.
 XX
 KW MAP1 gene; major antigenic protein 1; rickettsia; heartworm;
 KW DNA vaccine; ss.
 XX
 XX Cowdria ruminantium.
 XX
 XX WO9816554-Al.
 XX
 XX 23-APR-1998.
 XX
 XX 17-OCT-1997; 97WO-US19044.
 XX
 XX 17-OCT-1996; 96US-0733230.
 XX
 XX (UYFL) UNIV FLORIDA.

(MAP2) gene of rickettsial pathogens. The nucleic acid vaccines can be driven by the human cytomegalovirus (HCMV) enhancer-promoter. Cowdria ruminantium genes designated map 2, lhworf3, 4lhworf1, 18lhworf1 and 3gdorf3 may be used in therapeutic and diagnostic applications. The polypeptides are useful for detecting antibodies associated with infection by a rickettsial pathogen whilst the polynucleotides may be used to detect the presence of rickettsial nucleic acids.

Sequence 864 BP; 307 A; 139 C; 149 G; 269 T; 0 other;
SQ
XX

Sequence 864 BP; 307 A; 139 C; 149 G; 269 T; 0 other;

Alignment Scores:

Assignment Scores:	
Pred. No.:	1.1e-33
Score:	383.50
Percent Similarity:	50.17%
Best Local Similarity:	34.36%
Query Match:	25.40%
DB:	21
Length:	86%
Matches:	100
Conservative:	46
Mismatches:	108
Indels:	37
Gaps:	12

US-09-846-808-1 (1-284) x AAC68699 (1-864)

Qy	13	ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu-----	28
Db	37	ATATCATTAGTGGTCATTTTACCTGGTGTCTCTTTTCTGATGTAATCAGGAAGACAGC	96
Qy	29	AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer	48
Db	97	AACCCACAGCAGTGTTTACATTAGCCCAAAATACATGCCAACTGCATCACAATTTTGGT	156
Qy	49	LysLeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp	68
Db	157	AAAAATGTCAAATCAAGAAGATTCAAAAAATACTCAAACGGTATTTGGTCTAAAAAAGAT	216
Qy	69	LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPhe	83
Db	217	-----TGGATGGCGGTTTAAAAACCACTCAGATTCTAGCACTACTAATTCT	261
Qy	84	AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe	98
Db	262	ACAATTTTACTGAAAAGACTATTCTTCAGATATGAAAACAATCCGTTTATAGTTTC	321
Qy	99	SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118
Db	322	GCTGGACCAATTGGGTACTCAATGAAT--GGACCAAGAATAGAGTTCGAAGTATCCTAT	378
Qy	119	GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis	136
Db	379	GAATCTTTGATGTAAAAAACCTAGGTGGCACTATATAAAC---ACGCACACATGTAC	435
Qy	137	PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal	156
Db	436	TGTGCTTTAGTACAGCAGCACACAAAATAGCACTAATGGCGCAGGATTAACATCATCTGTT	495
Qy	157	ThrLeuIleAsnGlnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal	176
Db	496	ATGCTAAAAAACGAAATTTAAACAAATATATCATTAATGTTAAATCGGTGTATGATATC	555
Qy	177	-----AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAsp	194
Db	556	ATGCTTGATGAATACCA-----GTTCTCCATATGATGTGCAGGTATTGGCACTGAC	609
Qy	195	ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla	214
Db	610	TTAGTGTGCACTAATTAATGCTACAAATCCTTAATATTATCTATCAAGGAAGCTAGGCATA	669
Qy	215	SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe	234
Db	670	AGTTACTCAATCAATTCGAAGCTTCTATCTTTATCGGTGGACATTCCTAGAGTTATA	729
Qy	235	GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro	254
Db	730	GGTAATCAATTTAAA-----GATATTGCTACCTTAAAAATA	765
Qy	255	ProAsnSerLysProHisValHisThrThr-----AlaLeuAlaMetLeuSerIleGly	272


```

Db 67 TCTGCTAAGGAAGAAGAAATACACAGTTGGAGTCTTGGACTGAAGCAAAATTTGGGAC 126
      |||||
Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPhaAsnIle---LysTyrAsnPro 89
      :|||
Db 127 GGAAGCGCAATATCCCAACCTCCCAACAGATGATTTCACTGCTCTCAAAATTTATTCATT 186
      :|||
Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePhePheGlyTyrTyrAsnLys 108
      |||||
Db 187 AAATATGAANAACACCCGTTTGGTTTGCAGGAGCTATTGGTTACTCAATGGAT--- 243
      |||||
Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
      |||||
Db 244 GTTCCCAAGAATAGAGCTTCAAGCTATCTTATGAACATTTGATGATAAAATCAAGGTAAC 303
      |||||
Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
      |||||
Db 304 AATTATAGAAT---GAAGCACATAGATATGCTCTATCCCATAC---TCAGCAGCA 357
      |||||
Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
      :|||
Db 358 GACATGAGTAGTGAAGTAAATTTGCTTCTTAAANAATGAAGGATTACTTGACATA 417
      :|||
Qy 167 SerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleThrTyr 186
      |||||
Db 418 TCATTTATGCTGAACGCATGCTATGACGTAGTAGGCGAAGGCATACCTTTTCTCCTTAT 477
      |||||
Qy 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPhe 206
      |||||
Db 478 ATATGGCAGGATGCTGCTACTGATTTAGTATCCATGTTTGAAGCTACAAATCCTAAATTT 537
      |||||
Qy 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
      |||||
Db 538 TCTTACCAAGGAAGTTAGTTTAACTACTCTATPAGCCCAAGACCTTCTGTGTTATT 597
      |||||
Qy 227 GluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrPro 246
      |||||
Db 598 GTGGGCACCTTTCATGAAGTAATAGGGAACGAATTTAGAGATATTCCTACTATAATACCT 657
      |||||
Qy 247 Cys-----AspTyrProSerProThrProProAsnSerLys 258
      :|||
Db 658 ACTGGATCAACACTTCCAGGAAGAAAGAACTACCCCT----- 693
      |||||
Qy 259 ProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlySerIle 278
      :|||
Db 694 -----GCAATAGTAATACTGGATGATGCCACTTTGGAATAGAACTT 735
      |||||
Qy 279 GlyIleLysPheIle 283
      |||||
Db 736 GGAGGAAGGTTTGTA 750
      |||||
RESULT 31
AAC68706
ID AAC68706 standard; DNA; 830 BP.
XX
AC AAC68706;
XX
DT 02-MAR-2001 (first entry)
XX
DE Ehrlichia chaffeensis VSA5 gene partial coding sequence.
XX
KW Ehrlichia chaffeensis; VSA5; variable surface antigen 5; MAP1;
KW major antigenic protein 1; antirickettsial; vaccine; gene therapy;
KW Rickettsia; Cowdria; Anaplasma; map2; lhworf3; 4hworf1; 18hworf1;
KW 3gdorf3; ds.
XX
OS Ehrlichia chaffeensis.
XX
PN WO2000065063-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10886.

```

```

XX
PR 22-APR-1999; 99US-0130725.
XX
PA (UYFL ) UNIV FLORIDA.
XX
PI Barbet AF, Bowie MV, Ganta RR, Burrigle MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmore WW, Allemen AR;
XX
WPI: 2000-679675/66.
DR P-PSDB; AAB36189.
XX
New polynucleotides useful as DNA vaccines for conferring immunity to
rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
-
XX
Claim 4: Page 41-42; 63pp; English.
XX
The present sequence shows a high degree of similarity to the major
antigenic protein 1 (MAP1) gene of Ehrlichia sp. The MAP1 gene may be
used in a vaccines to protect animals or humans against rickettsial
diseases caused by a organisms of Rickettsia sp., Ehrlichia sp.,
Anaplasma sp. or Cowdria sp. The vaccine elicits an immune response
protective against the rickettsial pathogen. The nucleic acid vaccines
can be driven by the human cytomegalovirus (HCMV) enhancer-promoter.
Cowdria ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1
and 3gdorf3 may be used in therapeutic and diagnostic applications. The
polypeptides are useful for detecting antibodies associated with
infection by a rickettsial pathogen whilst the polynucleotides may be
used to detect the presence of rickettsial nucleic acids.
XX
SQ Sequence 830 BP; 278 A; 141 C; 160 G; 251 T; 0 other;
Alignment Scores:
Pred. No.: 2,36e-32 Length: 830
Score: 371.50 Matches: 94
Percent Similarity: 48.75% Conservative: 43
Best Local Similarity: 33.45% Mismatches: 113
Query Match: 24.60% Indels: 31
DB: 21 Gaps: 9
US-09-846-808-1 (1-284) x AAC68706 (1-830)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
      :|||
Db 39 ATATCATTAATATCTCTCTACCTGGAGTATCATTTCCGACCAGCAGCTAGTGGTATT 98
      :|||
Qy 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
      :|||
Db 99 AACGGTAATTTCTACATCAGTGGAAATACATCCCAAGTCTTCGCAATTTTGGAGTATTC 158
      :|||
Qy 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
      :|||
Db 159 TCTGCTAAGGAAGAAAGAAATACACAGTTGGAGTGTTCGACTGAAGCAAAATTTGGGAC 218
      :|||
Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPhaAsnIle---LysTyrAsnPro 89
      :|||
Db 219 GGAAGCGCAATATCCCAACCTCCCAACAGATGATTTCACTGCTCTCAAAATTTATTCATT 278
      :|||
Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
      :|||
Db 279 AAATATGAANAACACCCGTTTGGTTTGCAGGAGCTATTGGTTACTCAATGGAT--- 335
      :|||
Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
      :|||
Db 336 GTTCCCAAGAATAGAGCTTGAAGTATCTTATGAACAAATTTGATGATAAAATCAAGGTAAC 395
      :|||
Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
      :|||
Db 396 AATTATAGAAT---GAAGCACATAGATATGCTCTATCCCATAC---TCAGCAGCA 449
      :|||
Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
      :|||

```

```
Db 450 GACATGAGTAGTGCAGTAATAAATTTTGTCTTTCTAAATAAGTAAAGGATTACTTGACATA 509
Qy 167 SerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThrTyr 186
   ||| : : : : ||| ||||| ||| |||
Db 510 TCATTTATGCTGAAGCATGCTATGACCTAGTAGCGGAGGCATACCTTTTCTCCTTAT 569
Qy 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLysPhe 206
   ||| ||| ||| ||| : :
Db 570 ATATGCGCAGGTATCGGTACTGATTAGTATCCATGTTTGAAGCTACAAATCCTAAAT 629
Qy 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
   ||||| ||||| ||||| ||||| : : : : : ||||| |||||
Db 630 TCTTACCAGGAAGTATAGGTTTAACTGCTATTAAGCCAGAGCTTCTGTGTTAT 689
Qy 227 GluGlyTyrThrHisGlyLeuPheGlyLysLysPheGlyLysIleProValAsnTyrPro 246
   ||||| : : : : ||| : : ||| |||||
Db 690 GGTGGCACTTCATAGGTAATAGGACGAATTTAGAGATATCCCTACTATACTACT 749
Qy 247 Cys-----AspTyrProSerProThrProProAsnSerLys 258
   : : |||||
Db 750 ACTGATCAACACTTGCAGAAAGAACTACCT----- 785
Qy 259 ProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlySerIle 278
   ||||| : : ||| : : : : : |||||
Db 786 -----GCAATAGTAATACTGGATGTATGCCACTTTTGGAAAGAAATG 827
Qy 279 Gly 279
   |||
Db 828 GGA 830
   |||
RESULT 32
AAx34745
ID AAX34745 standard; DNA; 843 BP.
XX
AC AAX34745;
XX
DT 05-JUL-1999 (first entry)
XX
DE DNA encoding OMP-1C protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
KW detection; dog; ss.
XX
OS Ehrlichia chaffeensis.
XX
PN WO9913720-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US19600.
XX
PR 19-SEP-1997; 97US-0059353.
XX
PA (OHIS ) UNIV OHIO STATE.
XX
PI Ohashi N, Rikihisa Y;
XX
DR WPI; 1999-254290/21.
DR P-PSDB; AAY06945.
XX
PT Novel outer membrane proteins from Ehrlichia chaffeensis and
PT Ehrlichia canis
XX
PS Disclosure; Fig 5A; 55pp; English.
XX
CC The invention provides isolated outer membrane proteins (OMP) from
CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
CC in AAY06943-958. The E. canis proteins form part of the P30 family and
CC consist of proteins shown in AAY06959-970. The proteins and genes are
CC used to detect E. chaffeensis in patients and E. canis in dogs.
XX
SQ Sequence 843 BP; 262 A; 145 C; 169 G; 267 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 2,41e-32 Length: 843
Score: 371.50 Matches: 103
Percent Similarity: 50.69% Conservative: 44
Best local Similarity: 35.52% Mismatches: 114
Query Match: 24.60% Indels: 29
DB: 20 Gaps: 13

US-09-846-808-1 (1-284) x AAX34745 (1-843)
Qy 7 ArgLysPheValLeuTrpValMetLeuIleLeu-----PheThrProHisIleSer 23
   : : ||||| : : ||| |||
Db 10 AAAAAATTTTATAACAACATGCAATGGCATGGCAATGCTCTTCTTACCTGGAATATTA 69
Qy 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
   : : : : : ||| : : : : : |||||
Db 70 CTTTCTGAACAGTACAAGATGACAGTGTGAGTGGCAATTTCTATATTAGTGGCAAGTAC 129
Qy 41 LysProAlaArgGlnHisSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
   ||| : : : ||| ||||| |||||
Db 130 ATGCCAAGTCTTCTCATTTTGGAGTTTCTCTGCCAAGAAGAAAAAATCCTACTGTC 189
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
   : : : : : ||||| ||| : : :
Db 190 GCGTTGTATGTTTGAACAAGAT---TGGAAACGGTGTAGTGCTTCAAGTCATGCTGAT 246
Qy 81 ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
   : : : : : ||| : : : : : |||||
Db 247 GCGGACTTTAATAACAAGGTTATTTCTTTAAATACGAAAAACAATCCATTTCTAGGTTT 306
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
   : : ||| ||||| ||||| |||||
Db 307 GCAGGAGCTATTGGTTATTCATGGGT---GGTCCAAGATAGAGTTTGAAGTGTCCCTAT 363
Qy 119 GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGlyLysHis 136
   ||||| ||| : : ||||| ||| : : : :
Db 364 GAAACATTTGACGTGAAAAATCAAGGTGTAATTACAAAAAT---GATGCTCACAGATAC 420
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrVal 156
   ||||| : : : : : ||| ||| |||
Db 421 TGTGCTTAGATCGTAAAGCAAGC-----AGCACTAATGCCACAGTACCTACGCTG 474
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
   ||| ||| ||| : : : : : ||| |||||
Db 475 CTACTAAAAAATGAAGGACTACTTGATATATCATTATGTTGAATGCATGCTATGACGTA 534
Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
   : : : : : ||| : : ||| ||| |||
Db 535 GTAAAGTGAAGGAATACCTTTCTCT-----CCTTACATATGTCAGGTGTTGTACCGAT 588
Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
   ||| ||| ||||| ||||| |||||
Db 589 TTAATATCCATGTTTGAAGCTATAAACCTTAAATTTCTTATCAAGGAAGATTAGGTTTG 648
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
   ||||| : : : : : ||||| : : : : : |||||
Db 649 AGTTACTCTATAAACCCAGAGCTTCTGCTTTGTTGTTGGACATTTTCATAAAGTTGCA 708
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSer-----Pro 252
   ||| : : : ||| |||
Db 709 GGTAAATGAATTCAGGACATTTCTACTCTTAAAGCGTTTGTACACCATCATCTCAGCT 768
Qy 253 ThrProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGly 272
   ||||| ||||| ||| |||||
Db 769 ACTCCA-----GACTTAGCAACAGTAACACTGAGTGTGTGT 804
Qy 273 TyrTyrGlyGlySerIleGlyIleLysPhe 282
   : : : : : ||||| : : |||
Db 805 CACTTTGGAGTAGAAGCTTTGGAGGAAGATT 834
   : : : : : ||||| : : |||

RESULT 33
AAV07177
```


ID AAV07177 standard; DNA; 842 BP.
AC AAV07177;
XX
DT 14-SEP-1998 (first entry)
XX
DE Ehrlichia chaffeensis MAP1 gene coding sequence.
XX
KW MAP1 gene; major antigenic protein 1; rickettsia; DNA vaccine; ss.
XX
OS Ehrlichia chaffeensis.
XX
PN WO9816554-Al.
XX
PD 23-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19044.
XX
PR 17-OCT-1996; 96US-0733230.
XX
XX (UYFL) UNIV FLORIDA.
PA Barbet AF, Burridge MJ, Ganta RR, Mahan SM, McGuire TC;
PI Nyika A, Rurangirwa FR;
PI
DR WPI; 1998-251232/22.
DR P-PSDB; AAW51089.
XX
XX Composition containing nucleic acid encoding rickettsial antigen -
PT useful for, e.g. stimulating protective immune response in humans or
PT animals
XX
PS Claim 4; Page 17-18; 39pp; English.
XX
CC This DNA molecule comprises the coding region of the major antigen
CC protein 1 gene (MAP1) of Ehrlichia chaffeensis. It codes for a 280
CC amino acid MAP1 protein (see AAW51089). A claimed composition
CC comprises a nucleic acid (see AAV07176-82) encoding a polypeptide
CC (see AAW51088-99) that elicits a protective immune response against a
CC rickettsial pathogen. The nucleic acid is used, in human or
CC veterinary medicine, in vaccines to protect against Rickettsia,
CC Ehrlichia, Anaplasma and Cowdria species. The nucleic acids are
CC also useful as probes to identify related sequences, e.g. for
CC identification of organisms and for diagnosing infection. Use of
CC nucleic acid vaccines avoids the problem of protein purification
CC associated with protein-based vaccines. The nucleic acid does not
CC replicate in the host but remains episomal and capable of
CC expressing polypeptide for at least 19 mth.
XX
SQ Sequence 842 BP; 282 A; 143 C; 166 G; 251 T; 0 other;

Alignment Scores:
Pred. No.: 3,56e-32 Length: 842
Score: 370.00 Matches: 95
Percent Similarity: 49.12% Conservative: 45
Best Local Similarity: 33.33% Mismatches: 113
Query Match: 24.50% Indels: 32
DB: 19 Gaps: 9

US-09-846-808-1 (1-284) x AAV07177 (1-842)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAla-----SerValLeuAsn 29
XX
DB 37 ATCATTAATATCTCTCTTACTCGAGTATCATTTTCGACCCCAAGCGAGTAGTGGTC 96
XX
QY 30 AspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
XX
DB 97 ATTAACGGTAATTTCTACATCAGTGGAAAAATACGATGCCAAGGCTTCGCATTTCGAGTA 156
XX
QY 50 LeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
XX
DB 157 TTCTCTGCTAAGGAAGAAAGAAATACACAGTGGAGTGTGGAGCTGAAGCAAAATGG 216
XX

QY 70 LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
XX
DB 217 GACGGAAAGCGCAATATCCCAACTCTCCCAAAAGAGTATTCACCTGCTCAAAATATTCA 276
XX
QY 89 ProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsn 107
XX
DB 277 TTTAAATATGAACAACACCCGTTTGTAGGTTTTCAGGAGCTATTGGTTACTCAATGGAT 336
XX
QY 108 LysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn----- 125
XX
DB 337 ---GGTCCAAGAATACAGCTTGAAGTATCTTATGAACATTTGATGTATAAAATCAAGGT 393
XX
QY 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGly 145
XX
DB 394 AACAATATAAAGAT---GAAGCACATAGATATGTGCTCTATCCCATAC---TCAGCA 447
XX
QY 146 GlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
XX
DB 448 GCAGACATGAGTAGTCAAGTAAATTTGTCTTCTTAAAAAATGAAGGATTACTTTGAC 507
XX
QY 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleThr 185
XX
DB 508 ATATCATTTATGCTGAACGCTATGAGTAGTAGCGAAGCATACCTTTTCTCCT 567
XX
QY 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
XX
DB 568 TATATATGCGAGGTATCGTACTGATTTAGTATCCATGTTTGAAGCTACAAATCCTAAA 627
XX
QY 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
XX
DB 628 ATTCTTACCAAGAAAGTGTAGGTTTAAGCTACTTATAGCCCAAGGCTTCGTGTTT 687
XX
QY 226 IleGluGlyTyrThrHisGlyLeuPheGlyLysPheGlyLysIleProValAsnTyr 245
XX
DB 688 ATTTGGTGGCATTTCATAGGTAATAGGGAAGCAATTTAGACATATTCCTACTATAATA 747
XX
QY 246 ProCys-----AspTyrProSerProThrProProAsnSer 257
XX
DB 748 CCTACTGGATCAACACTTGCAGGAAAAAGAACTACCT----- 786
XX
QY 258 LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySer 277
XX
DB 787 -----GCATAGTAATACTGGATGTATGCCACTTTGGAATAGAA 825
XX
QY 278 IleGlyIleLysPhe 282
XX
DB 826 ATGGGAGGAGGTTT 840
XX
RESULT 34
AAC68700
ID AAC68700 standard; DNA; 842 BP.
XX
AC AAC68700;
XX
XX 02-MAR-2001 (first entry)
DT
XX
DE Ehrlichia chaffeensis MAP1 gene.
XX
KW Ehrlichia chaffeensis; MAP1; major antigenic protein 1; antirickettsial;
KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; lhworf3;
KW 4hworfl; 18hworfl; 3gdorf3; ds.
OS
XX Ehrlichia chaffeensis.
XX
PN WO2000065063-A2.
XX
PD 02-NOV-2000.
XX
XX 21-APR-2000; 2000WO-US10886.
PF
XX 22-APR-1999; 99US-0130725.
PR
XX (UYFL) UNIV FLORIDA.
PA

XX
PI Barbet AF, Bowie MV, Ganta RR, Burrige MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmire WW, Alleman AR;
XX
DR WPI; 2000-679675/66.
DR P-PSDB; AAB36183.

Alignment Scores:		
Pred. No.:	3.56E-32	Length:
Score:	370.00	Matches:
Percent Similarity:	49.12%	Conservative:
Best Local Similarity:	33.33%	Mismatches:
Query Match:	24.50%	Indels:
DB:	21	Gaps:
		9
US-09-846-808-1 (1-284)	x AAC68700 (1-842)	

CC polypeptides are useful as vaccines for conferring immunity to rickettsia
 CC infection, including Cowdria ruminantium causing heartwater. The MAP
 CC polynucleotides may be used as molecular markers in nucleic acid
 CC analysis procedures, and to produce the MAP polypeptides, which may
 CC be used to raise antibodies that are reactive with the polypeptides.
 CC The nucleic acids may further be used as probes to identify
 CC complementary sequences within other nucleic acid molecules or genomes,
 CC where such probes can be applied to identify or distinguish infectious
 CC strains of organisms in diagnostic procedures or in rickettsial
 CC research where identification of particular organisms or strains is
 CC needed.

XX Sequence 842 BP; 282 A; 143 C; 166 G; 251 T; 0 other;

Alignment Scores:

Pred. No.: 3,56e-32 Length: 842
 Score: 370.00 Matches: 95
 Percent Similarity: 49.12% Conservative: 45
 Best Local Similarity: 33.33% Mismatches: 113
 Query Match: 24.50% Indels: 32
 DB: 22 Gaps: 9

US-09-846-808-1 (1-284) x AAS07576 (1-842)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAla-----SerValLeuAsn 29
 DB 37 ATCATTAATATCTCTCTTACCTGGAGTATCATTTTCGCCCAAGCAGGTAGTGTC 96
 QY 30 AspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB 97 ATTAACGGTAATTTTACATCAGTGGAAATACGATGCCAAGGCTTCGCATTTTGGAGPA 156
 QY 50 LeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
 DB 157 TTCTCTGCTAAGCAAGAAATACACAGTGGNGTCTTTGGACTGAAGCAAAATTCG 216
 QY 70 LeuAsnAspLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
 DB 217 GACGGAAGCGCAATATCCAACTCTCCCAACAGATGATTCACCTGTCTCAAATATTCA 276
 QY 89 ProTyrTyrGluAsnArg---LeuGlyPheSerGlyLePheGlyTyrTyrAsn 107
 DB 277 TTTAATATGAACAACCCGTTTGTAGGTTTTCGAGGAGCTATTGGTTACTCAATGGAT 336
 QY 108 LysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn----- 125
 DB 337 ---GGTCCAAGATAGAGCTTGAAGTATCTTATGAACATTCATGATAAAATCAAGCT 393
 QY 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluLeSerGly 145
 DB 394 AACAAATTATAAAGAT---GAAGCACATAGATATTGTCTCTATCCATAAC---TCAGCA 447
 QY 146 GlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
 DB 448 GCAGACATCAGTAGTCAAGTAAATATTTGTCTTTCTAAATAAAGAGGATTACTTTGAC 507
 QY 166 ThrSerAlaLeuLeuAsnValCysTyrAspValAspValGlyLeuLysHisAsnIleThr 185
 DB 508 ATATCATTTATGCTGAACCATGCTATGATGAGTAGGAGGAGGATACCTTTTCTCTCT 567
 QY 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
 DB 568 TATATATGCGCAGGTATCGTACTGATTTAGTATCCATGTTTGAAGCTACAAATCCTAAA 627
 QY 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
 DB 628 ATTCTTACCAAGGAAGTTAGTTTAAGCTACTCTATAGGCCCAAGACTCTCTGTGTT 587
 QY 226 IleGluGlyTyrTyrHisGlyPheGlyLysLysPheGlyLysIleProValAsnTyr 245
 DB 688 ATTTGGTGGCCACTTTCATAGATTAAGGAACGAATTTAGAGATATTCCTACTATAATA 747
 QY 246 ProCys-----AspTyrProSerProThrProAsnSer 257

DB 748 CCTACTGGATCAACACTTCGAGGAAAGAACTACCCCT----- 786
 QY 258 LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlyser 277
 DB 787 -----GCAATAGTAATACTGGATGATGCCACITTTGGAATAGAA 825
 QY 278 IleGlyIleLysPhe 282
 DB 826 ATGGGAGGAAGGTTT 840
 RESULT 36
 AAX34761
 ID AAX34761 standard; DNA; 924 BP.
 XX
 AC AAX34761;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE DNA encoding p30-1 protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
 KW detection; dog; ss.
 XX
 OS Ehrlichia canis.
 XX
 PN WO9913720-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19600.
 XX
 PR 19-SEP-1997; 97US-0059353.
 XX
 PA (OHIS) UNIV OHIO STATE.
 XX
 PI Ohashi N, Rikihisa Y;
 XX
 DR WPI; 1999-254290/21.
 DR P-PSDB; AAY06961.
 XX
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 PS Disclosure; Fig 21A; 55pp; English.
 XX
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the p30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 SQ Sequence 924 BP; 301 A; 158 C; 160 G; 305 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,52e-31 Length: 924
 Score: 363.00 Matches: 94
 Percent Similarity: 50.18% Conservative: 45
 Best Local Similarity: 33.94% Mismatches: 118
 Query Match: 24.04% Indels: 20
 DB: 20 Gaps: 9
 US-09-846-808-1 (1-284) x AAX34761 (1-924)
 QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
 DB 124 ATATCATTTATGACTCTATTCACAGCATATCTTTCTGATACTATACAGATGGTAAC 183
 QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB 184 ATGGGTGGTAACCTTCTATATTAGTGGAAAGATGTACCAAGTCTCTACACATTTTGTAGC 243

```
Qy 50 LeuLeuLeuLeuGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
Db 244 TTCACAGTAAGAAGAAAGCAACTGTTGGAGTTTTTGGATTAAACATGATTGG 303
Qy 70 LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
Db 304 GATGGAAGTCCAATACTTAAGAATAAA---CACGGTGACTTACTGTTCACAACTATTTCG 360
Qy 89 ProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsn 107
Db 361 TTCAGATACAGAGAACAATCCATTCTAGGTTTCGAGGAGCTATCGGTACTCAATGGGT 420
Qy 108 LysAsnPheArgIleGluSerGlyLeuSerTyrGluThrPheHisIleLys-----Asn 125
Db 421 ---GGCCCAAGATAGATTGCAATATCTTATGAGCATTCGAGCTAAAAGTCCCTAAT 477
Qy 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGly 145
Db 478 ATCAATTATCAAAAT---GACGCGCACAGGTACTGCGCTCTATCTCATCACATCGGCA 534
Qy 146 GlySerAsnAsnProAlaAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
Db 535 GCCATGGAA-----GCTGATAAATTTGCTCTTAAATAAAGCAAGGGTTAATTGAC 585
Qy 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThr 185
Db 586 ATATCACTTGCAATAAATGCATGTATGATATATAATAATGACAAAGTACCTGTTCTCCT 645
Qy 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
Db 646 TATATATCGCAGGTATTGGTACTGATTTGATTCTTATGTTGAAGCTACAAAGTCCCTAAA 705
Qy 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
Db 706 ATTCTTACCAGGAATACTGGCATTTAGTTACTCTATTAATCGGAACCTCTGTTTTC 765
Qy 226 IleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysIleProValAsnTyr 245
Db 766 ATCGTGGGCATTTCCACAGCATATAGTAAATGAGTTTAGAGATATTCCTGCAATAGTA 825
Qy 246 ProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrAla 265
Db 826 CCTAGTAACCTCAACTACAATAAAGTGGACCAAA-----TTTGCA 864
Qy 266 LeuAlaMetLeuSerIleGlyTyrGlyGlySerIleGlyIleLysPhe 282
Db 865 ACAGTAACACTAAATATGTGTGCTTTGGTTTAGAACTTTGGAGGAAGATT 915

RESULT 37
AAD01292
ID AAD01292 standard; DNA; 1607 BP.
XX
AC AAD01292;
XX
DT 12-OCT-2000 (first entry)
XX
DE Ehrlichia canis immunoreactive protein Eca28-1 DNA.
XX
KW Homologous 28-kDa protein gene; Eca28-1; immunoreactive; vaccine;
KW p28 gene; polymorphic multiple gene family; immunoprotective antigen;
KW antibacterial; canine ehrlichiosis; canine tropical pancytopenia;
KW tick-borne rickettsial disease; serodiagnosis; ds.
XX
OS Ehrlichia canis.
XX
FH Key Location/Qualifiers
FT CDS 146..982
FT FT /*tag= a
FT FT /product= "Eca28-1 protein (30-kDa)"
FT FT 146..214
FT FT /*tag= b
FT FT 215..979
FT FT /*tag= c
```

```
FT XX /product= "Mature Eca28-1 protein (28-kDa)"
PN WO200032745-A2.
XX 08-JUN-2000.
XX 24-NOV-1999; 99WO-US28075.
XX 30-NOV-1998; 98US-0201458.
XX 03-MAR-1999; 99US-0261358.
XX (RERE-) RES DEV FOUND.
XX Walker DH, Yu X, McBride JW;
PI WPI; 2000-412298/35.
XX P-PSDB; AAY71477.
XX Ehrlichia canis antigens useful for vaccinating against canine
PT ehrlichiosis in dogs
XX Claim 5; Fig 1; 86pp; English.
XX The patent relates to homologous 28-kilodalton (kDa) protein genes of
CC Ehrlichia canis, designated Eca28SA1, Eca28SA2, Eca28SA3, Eca28-1 and
CC Eca28-2. These genes are members of a polymorphic multiple gene family
CC and contained in a single locus of 5.592 kb. The 28-kDa proteins are
CC immunoreactive with anti-E. canis serum hence are important
CC immunoprotective antigens. The protein is useful for vaccinating
CC against E. canis infections such as canine ehrlichiosis in dogs.
CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
CC tick-borne rickettsial disease of dogs. Eca28-1 is conserved amongst
CC different strains of E. canis and hence useful for serodiagnosis of
CC canine ehrlichiosis. The present sequence is a DNA encoding E. canis
CC Eca28-1 30-kDa protein which is post-translationally modified to a
XX mature 28-kDa protein by cleavage of N-terminal signal sequence.
SQ Sequence 1607 BP; 543 A; 268 C; 250 G; 540 T; 6 other;

Alignment Scores:
Pred. NO.: 5.59e-31 Length: 1607
Score: 363.00 Matches: 94
Percent Similarity: 50.18% Conservative: 45
Best Local Similarity: 33.94% Mismatches: 118
Query Match: 24.04% Indels: 20
DB: 21 Gaps: 9

US-09-846-808-1 (1-284) x AAD01292 (1-1607)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 182 ATATCATTAATGTTACTCTATTCACAGCATATCTTTCTGTACTATATACAGATGTTAAC 241
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
Db 242 ATGGGTGGTAACCTCTATATTAGTGGAAAGTATGTACCAAGTCTCACATTTTGGTAGC 301
Qy 50 LeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
Db 302 TTCTCAGCTAAGAAGAAAGCAATCAACTGTTGGAGTTTTTGGATTAAACATGATTGG 361
Qy 70 LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
Db 362 GATGGAAGTCCAAATACTTAAGAATAAA---CACGGTACTTTTACTTCCCAACTATTTCG 418
Qy 89 ProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsn 107
Db 419 TTCAGATACAGAGAACAATCCATTTCTAGGTTTTCGAGGAGCTATCGGTACTCAATGGGT 478
Qy 108 LysAsnPheArgIleGluSerGlyLeuSerTyrGluThrPheHisIleLys-----Asn 125
Db 479 ---GGCCCAAGAATAAGATTGCAATATCTTATGAAGCATTCGAGGTAATAAGTCCCTAAT 535
```

```
QY 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGly 145
Db 536 ATCAATATCAAAAT---GACGCGCAGGTAAGTGGCTCTATCATCACACATCGGCA 592
QY 146 GlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
Db 593 GCATGGAA-----GCTGATAAATTTGCTCTTTAAACAAAGGTTAATTGAC 643
QY 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThr 185
Db 644 ATATCACTTGCATAAATCATGTTATGATATAATAAATGACAAAGTACTGTTCTCCT 703
QY 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLys 205
Db 704 TATATATGCCAGGTATTGCTACTGATTGATTCTATGTTTGAAGCTACAAGCTCTAAA 763
QY 206 PheSerTyrGlnCysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
Db 764 ATTTCTTACCAAGGAACCTGGGCATAGTTACTTCTATTATCCGGAACCTCTGTTTC 823
QY 226 IleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysProValAsnTyr 245
Db 824 ATCGGTGGCCATTCCACAGGATCATAGTAATGAGTTAGAGATATCTCTCAATAGTA 883
QY 246 ProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrAla 265
Db 884 CCTAGTAACCTCAACTACAATAAGTGGACCAAA-----TTTGCA 922
QY 266 LeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPhe 282
Db 923 ACAGTAACCAATAATGTGTGCTACTTGTGTTAGAACCTGGGAGGAGATT 973
RESULT 38
AAV07180
ID AAV07180 standard; DNA; 1570 BP.
XX AC AAV07180;
XX DT 14-SEP-1998 (first entry)
XX DE Ehrlichia canis VSA genomic locus.
XX KW MAP1 homologue; variable surface antigen; VSA1; VSA2; rickettsia;
XX OS Ehrlichia canis.
XX FH Key Location/Qualifiers
FT RBS 1..5 /*tag= a
FT CDS 11..874 /*tag= b
FT /*note= "VSA1 gene"
FT terminator 1015..1022 /*tag= c
FT terminator 1027..1034 /*tag= d
FT -35_signal 1081..1086 /*tag= e
FT exon 1101..1106 /*tag= f
FT RBS 1161..1165 /*tag= g
FT CDS 1171..1560 /*tag= h
FT /*note= "truncated VSA2 gene"
PN W09816554-A1.
XX 23-APR-1998.
XX PF 17-OCT-1997; 97WO-US19044.
XX
```

```
PR 17-OCT-1996; 960S-0733230.
XX (UYFL ) UNIV FLORIDA.
XX Barbet AF, Burr ridge MJ, Ganta RR, Mahan SM, McGuire TC;
PI Nyika A, Rurangirwa FR;
XX WPI; 1998-251232/22.
DR P-PSDB; AAW51096-97.
XX Composition containing nucleic acid encoding rickettsial antigen -
PT useful for, e.g. stimulating protective immune response in humans or
PT animals
XX Claim 4; Fig 2C; 39pp; English.
XX This is the DNA sequence of a 1.5 genomic locus of Ehrlichia canis
CC that was obtained using a PCR cloning strategy based on identifying
CC genes homologous to the major antigenic protein MAP1 (see AAW51088)
CC of Cowdria ruminantium. It includes 2 very similar but
CC non-identical open reading frames (ORFs), of which ORF2 is a
CC partial gene. Due to their similarity to MAP1 surface antigen
CC genes of C. ruminantium, the E. canis ORFs are designated variable
CC surface antigen (VSA) genes 1-2. A claimed composition comprises a
CC nucleic acid (see AAV07176-82) encoding a polypeptide (see AAW51088-99)
CC that elicits a protective immune response against a rickettsial
CC pathogen. The nucleic acid is used, in human or veterinary
CC medicine, in vaccines to protect against Rickettsia, Ehrlichia,
CC Anaplasma and Cowdria species. The nucleic acids are also useful
CC as probes to identify related sequences, e.g. for identification of
CC organisms and for diagnosing infection. Use of nucleic acid
CC vaccines avoids the problem of protein purification associated with
CC protein-based vaccines. The nucleic acid does not replicate in the
CC host but remains episomal and capable of expressing polypeptide for
CC at least 19 mth.
XX SQ Sequence 1570 BP; 565 A; 223 C; 251 G; 531 T; 0 other;
Alignment Scores:
Pred. No.: 1,12e-28 Length: 1570
Score: 342.50 Matches: 95
Percent Similarity: 45.70% Conservative: 43
Best Local Similarity: 31.46% Mismatches: 129
Query Match: 22.68% Indels: 35
DB: 19 Gaps: 9
US-09-846-808-1 (1-284) x AAV07180 (1-1570)
QY 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIleLeu----- 17
Db 5 GTAAATATGAAATATAAAAAAACTTTTACAGTAACATGTCATTATTATTAACCTCTTT 64
QY 18 -----PheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyr 35
Db 65 ACACATTTTATACCTTTTATAGTCAGCAGCGTCACATTTTGGAAATTTTTCAGCTAAGAAGAA 181
QY 36 ValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuIleLysGluSer 55
Db 122 ATTAGTGGAAATATATATGATCCACAGCGTCACATTTTGGAAATTTTTCAGCTAAGAAGAA 181
QY 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu----- 70
Db 182 CAAAGTCTTTACTAAGGTATTAGTTGGTTAGATCAACGATTATACATAATATTATAAAC 241
QY 71 ---AsnAspLeuLeuThrGlyIleLys---AspAsnThrAsnPheAsnIleLysTyrAsn 88
Db 242 AATAATGATACAGCAAGAGCTCTTAAGGTTCAAAATTTTCAATTAATAACAAATAAAC 301
QY 89 ProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
Db 302 CCATTT-----CTAGGATTGTCAGGAGCTATTGGTTATTCAATAGGC--- 343
QY 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsnAsnGlyTyr 128
```

```
Db 344 AATTCAAGATAGACTAGAGTATACATGAATATATTGTACTATAAAACCCAGGAAC 403
Qy 129 Lysargille---AspCysGLuLysHisPheAlaLeuAlaLysGluIle----- 143
Db 404 AATTATTAAATGACTCTCACAAATATTGCCGTTATCTCATCGAAGTCACATATGCAGT 463
Qy 144 -----SerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuLeasn 160
Db 464 GATGAAATAGCGGAGATTGTTACACTGCAAAACTGATAAGTTGTACTTCTGAAAAAT 523
Qy 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys 180
Db 524 GAAGTTTACTTGAGCTCTATTATGTTAAACGATGTTATGACATACAACTGAAAAA 583
Qy 181 HisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
Db 584 ATGCCCTTTTACCTATATATGTCAGGTATTGGTACTGATCTCATATCTATGTTGAG 643
Qy 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
Db 644 ACAACACAAAAAATATCTTATCAAGGAAAGTTAGGTTTAAACTATATAAATCA 703
Qy 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys 240
Db 704 AGAGTTTCTGTTTTCAGGTTGGGCACCTTCATTAAGGTAATAGGTAATGAATTTAAAGGT 763
Qy 241 IleProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHis 260
Db 764 ATT-----CCCTACTCTATTACCTGATGGTCAACAACTTAAA 799
Qy 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrGlyGlySerIleGlyIle 280
Db 800 GTACAACAGTGTGCAACAGTAACTAGATGTTGCCATTTTCGGTTAGAGATTGGAAGT 859
Qy 281 Lysphe 282
Db 860 AGATTT 865

RESULT 39
ID AAS07583 standard; DNA; 1570 BP.
XX
AC AAS07583;
XX
XX
XX 23-OCT-2001 (first entry)
XX
XX DNA encoding variable surface antigens 1-2 (VSA1-2) from E. canis.
XX
XX Major antigenic protein; MAP; vaccine; immunogenic; rickettsia;
KW infection; heartwater; diagnostic; variable surface antigen: VSA; ds.
XX
XX Ehrlichia canis.
XX
XX key Location/Qualifiers
XX RBS 1..5
XX FT /*tag= a
XX FT /note= "Ribosome binding site of VSA1 gene"
XX FT 11..874
XX FT /*tag= b
XX FT /product= "Variable surface antigen 1 (VSA1)"
XX FT 1015..1020
XX FT terminator
XX FT /*tag= c
XX FT /note= "Transcription terminator of VSA1 gene"
XX FT 1027..1034
XX FT /*tag= d
XX FT /note= "Transcription terminator of VSA1 gene"
XX FT 1081..1086
XX FT -35_signal
XX FT /*tag= e
XX FT /note= "-35 region of VSA2 gene"
XX FT 1101..1106
XX FT -10_signal
XX FT /*tag= f
XX FT /note= "-10 region of VSA2 gene"
```

```
FT RBS 1151..1155
FT FT /*tag= g
FT FT /note= "Ribosome binding site of VSA2 gene"
FT CDS 1161..1169
FT FT /*tag= h
FT FT /product= "variable surface antigen 2 (VSA2)"
FT FT /partial
FT FT /note= "No stop codon"
FN US6251872-B1.
XX
XX 26-JUN-2001.
XX
XX 17-OCT-1997; 97US-09533326.
XX
XX 17-OCT-1996; 96US-0733230.
XX (UYFL ) UNIV FLORIDA.
XX
XX Barbet AF, Ganta RR, McGuire TC, Burridge MJ, Nyika A;
XX Rurangirwa FR, Mahan SM, Bowie MV, Alleman AR;
XX
XX WPI: 2001-424487/45.
XX P-PSDB; AAU04200, AAU04201.
XX
XX New MAP2 genes and polypeptides useful as vaccines for conferring
XX immunity to human and animal rickettsial diseases, e.g. heartwater, or
XX as molecular markers in nucleic acid analysis procedures -
XX
XX Example 3; Fig 2C; 30pp; English.
XX
XX The sequence represents the coding sequence of variable surface antigen
XX (VSA) gene locus encoding VSA1-2 proteins of Ehrlichia canis, which
XX have similarity to major antigen proteins (MAP). The MAP polynucleotides
XX and polypeptides are useful as vaccines for conferring immunity to
XX rickettsia infection, including Cowdria ruminantium causing heartwater.
XX The MAP polynucleotides may be used as molecular markers in nucleic acid
XX analysis procedures, and to produce the MAP polypeptides, which may
XX be used to raise antibodies that are reactive with the polypeptides.
XX The nucleic acids may further be used as probes to identify
XX complementary sequences within other nucleic acid molecules or genomes,
XX where such probes can be applied to identify or distinguish infectious
XX strains of organisms in diagnostic procedures or in rickettsial
XX research where identification of particular organisms or strains is
XX needed.
XX
XX SQ Sequence 1570 BP; 565 A; 223 C; 251 G; 531 T; 0 other;

Alignment Scores:
Pred. NO.: 1.12e-28 Length: 1570
Score: 342.50 Matches: 95
Percent Similarity: 45.70% Conservative: 43
Best Local Similarity: 31.46% Mismatches: 129
Query Match: 22.68% Indels: 35
DB: 22 Gaps: 9

US-09-846-808-1 (1-284) x AAS07583 (1-1570)
Qy 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIleLeu----- 17
Db 5 GTAAATATGAAATATAAAAAAATTTTACAGTAACTGCATTAGTATTATTACTTCCTTT 64
Qy 18 -----PheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyr 35
Db 65 ACACATTTTATACCTTTTATAGTCCACGCGTCCAGTACATTCACAAAC---TTCAC 121
Qy 36 ValGlyIleGlnTyrLysProAlaArgGlnHisSerLysLeuLeuIleLysGluSer 55
Db 122 ATTAGTGGAATAATATATGCCAACGCGTCACATTTTGGAAATTTTTCAGCTAAAGAAGAA 181
Qy 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu----- 70
Db 182 CAAAGTTTCTACTAAGGTATTAGTGGGTTAGATCAACGATTCATCATCAATAATATTATAAC 241
```

```

QY 71 ---AsnAspLeuLeuThrGlyIleLys---AspAsnThrAsnPheAsnIleLysTyrAsn 88
DB 242 AATAATGATACAGCAAGAGCTCTAAGGTTCAAAATTAATCATTTAAATACAAAATAAC 301
QY 89 ProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
DB 302 CCATTT-----CTAGGATTGCGAGGCTATTGGTTATTCAATAGGC--- 343
QY 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsnAsnGlyTyr 128
DB 344 AATTCAGAAATAGCACTAGCAATGATCATGAAATATTGATCTAATAACCCAGGAAC 403
QY 129 LysArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIle----- 143
DB 404 AATTATTAAATGACTCTCACAATAATTGCGCTTATCTCATGGAAGTACATATGCAGT 463
QY 144 -----SerGlyCysSerAsnProAlaAsnLysTyrValThrLeuIleAsn 160
DB 464 GATGGAATAGCGGAGATTGGTACACTGCAAAACCTGATAAGTTTGTCTTCTGAAAAAT 523
QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys 180
DB 524 GAAGGTTTACTTCACGCTCTCATTTATGTTAAACGCATGTTATGACATACAACTGAAAA 583
QY 181 HisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
DB 584 ATGCCCTTTTTCACCTTATATATGTCAGGCTATTGGTACTGATCTCATATCTATGTTGAG 643
QY 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
DB 644 ACAACAACAAACAAATATCTTATCAAGGAAAGTTAGGTTTAAACTATATCTATAAACTCA 703
QY 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys 240
DB 704 AGAGTTTCTGTTTTCAGGTGGCACTTTCATAGGTAATAGGTAATCAATTAAGCT 763
QY 241 IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis 260
DB 764 ATT-----CCTACTCTATTACCTGATGGATCAACATTAATAA 799
QY 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyCysIleGlyIle 280
DB 800 GTACAACAGCTCTGCAACAGTACATAGATGTGTGCCATTTCGGGTTAGAGATTGGAAGT 859
QY 281 LysPhe 282
DB 860 AGATTT 865
RESULT 40
ID AAX34764
XX AAX34764 standard; DNA; 831 BP.
AC AAX34764;
XX 05-JUL-1999 (first entry)
XX DNA encoding p30-4 protein.
XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
XX detection; dog; ss.
OS Ehrlichia canis.
PN WO9913720-A1.
XX 25-MAR-1999.
XX 18-SEP-1998; 98WO-US19600.
XX 19-SEP-1997; 97US-0059353.
XX (OHIS ) UNIV OHIO STATE.

```

```

XX Ohashi N, Rikihisa Y;
XX WPI: 1999-254290/21.
XX P-PSDB; AAY06964.
XX Novel outer membrane proteins from Ehrlichia chaffeensis and
XX Ehrlichia canis
XX Disclosure; Fig 24A; 55pp; English.
XX The invention provides isolated outer membrane proteins (OMP) from
XX Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
XX of the OMP family and consist of proteins OMP-1, -(B to Z) shown
XX in AAY06943-958. The E. canis proteins form part of the p30 family and
XX consist of proteins shown in AAY06959-970. The proteins and genes are
XX used to detect E. chaffeensis in patients and E. canis in dogs.
XX Sequence 831 BP; 300 A; 126 C; 136 G; 269 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 5.82e-29 Length: 831
Score: 341.50 Matches: 90
Percent Similarity: 50.70% Conservative: 54
Best Local Similarity: 31.69% Mismatches: 117
Query Match: 22.62% Indels: 23
DB: 20 Gaps: 10
US-09-846-808-1 (1-284) x AAX34764 (1-831)
QY 7 ArgLysPheValLeuTrpValMetLeuIleLeuPheThr-----ProHisIleSer 23
DB 10 AAAAAATTTCTTATACAACTACATTTGGTATCTACTAACAACTCTTTTACCTGCATATCT 69
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
DB 70 TTTCTCCAAACCAATACATGAAACAACTACTACAGGAACCTTTTACATATTATGGAANAATAT 129
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
DB 130 GTACCAAGTATTTCACATTTTGGGAACCTTTTTCAGCTAAAGAGAAAAACAACTACT 189
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
DB 190 GGAATTTTGGATTAAAGAAATCATGCTGCTGGTATCATCTCTGAT---AAGAACAT 246
QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
DB 247 GCAGCTTTTAATATCCCAATATTATTCATTAAATATGAAATATTAATCCATTTTAGGATTT 306
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
DB 307 CGAGGGGTAATTGGCTATTCAATAGTAGTATGCCA---AGAATAGCAATTTGAAGTATCATAC 363
QY 119 GluThrPheHisIleLysAsnAsnGlyTyrLys---ArgIleAspCysGluLysHisPhe 137
DB 364 GAGACATTTCGATGTACAAAATCCAGAGATAAGTTTAAACAATGATGCACATAAGTATTTGT 423
QY 138 AlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrValThr 157
DB 424 GCTTATCCCAATGAT-----TCCAGTAAACCAATGAAAGCTGTAATTCGTTTTT 474
QY 158 LeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAsp 177
DB 475 CTCAAAAATGAAGGATTAAAGTACATATCACATCATCTGTTAAATCTATGTTATGATATAATA 534
QY 178 GlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp 197
DB 535 AACAAAAAGAATGCCTTTTTCACCTTACATATGTGCGAGCATTTGCTACTGACTTAATATTC 594
QY 198 PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr 217
DB 595 ATGTTTGACGCTATATAAACCATCAAGCTGCTTATCAAGGAAATTAGGTTTAAATATCA 654

```

Qy	218	ValSerProGlnValSerValPheIleGluGlyTyrTyHisGlyLeuPheGlyLysLys	237
Db	655	ATAAGCCCAAGACTAACTTTCTATGGGTGTCACATTTTCACAAAGTAAACAACAACGAG	714
Qy	238	PheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProProAsnSer	257
Db	715	TTT---AGAGTTCTGTCTATTAACTGCTGGAGGACTCGCTCCAGAT-----	759
Qy	258	LysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyGlyGlySer	277
Db	760	-----AATCTATTGGCAATAGTAAGACTGAGTATATGTCATTTTGGGTAGAA	807
Qy	278	IleGlyIleLys	281
Db	808	TTTGGGTACAGG	819

Search completed: October 19, 2002, 03:57:49
Job time : 215 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 19, 2002, 03:00:25 ; Search time 1872 Seconds
(without alignments)
3174.752 Million cell updates/sec

Title: US-09-846-808-1
Perfect score: 1510
Sequence: 1 MSRRNRKFWLWMLILFTP.....ALAMLSIGYGGSIGIKFIL 284

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-O=/cgn2_1/USPRO_spool/US09846808/runat_18102002_140435_26016/app_query.fasta_1.455
-DB=genEmbl -QWRT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09846808 -CGN_1_1.1182 @runat_18102002_140435_26016 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_em:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1510	100.0	14759	1	AF230642 Ehrlichia
2	1510	100.0	27190	1	ECU72291 Ehrlichia c
3	932	61.7	28254	1	AF078553 Ehrlichia
4	448	29.7	11329	1	AF082744 Ehrlichia
5	425	28.1	6913	1	AF324792 Ehrlichia
6	424	28.1	3535	1	AF125274 Ehrlichia r
7	424	28.1	3538	1	AF125277 Cowdria r
8	424	28.1	3541	1	AF125279 Cowdria r
9	424	28.1	3572	1	AF125278 Cowdria r
10	422	27.9	3507	1	AF125276 Cowdria r
11	422	27.9	3551	1	AF125275 Cowdria r
12	411.5	27.3	843	6	AX042314 Sequence
13	411.5	27.3	4683	1	AF062761 Ehrlichia
14	397.5	26.3	1263	1	CRU50830 Cowdria rum
15	393	26.0	1101	1	CRU50834 Cowdria rum
16	392	26.0	1278	1	CRU50832 Cowdria rum
17	391	25.9	1564	1	CRU49843 Cowdria rum
18	389	25.8	873	1	AF368001 Cowdria r
19	389	25.8	828	1	AF368007 Cowdria r
20	388	25.7	1278	1	AF368014 Cowdria r
21	388	25.7	1467	1	CRU50835 Cowdria rum
22	387.5	25.7	837	6	AX042313 Sequence
23	386	25.6	1283	1	AF077732 Ehrlichia
24	386	25.6	1307	1	AF393394 Ehrlichia
25	386	25.6	1309	1	AF393390 Ehrlichia
26	385	25.5	1243	1	AF077733 Ehrlichia
27	385	25.4	807	1	AF368004 Cowdria r
28	384	25.4	2977	1	AF068234 Ehrlichia c
29	384	25.4	27190	1	ECU72291 Ehrlichia c
30	384	25.4	864	6	AX042305 Sequence
31	383.5	25.4	1265	1	CRU50833 Cowdria rum
32	383.5	25.4	28254	1	AF078553 Ehrlichia
33	382	25.3	831	1	AF368000 Cowdria r
34	381.5	25.3	816	1	AF325176 Cowdria s
35	381.5	25.3	816	1	AF368013 Cowdria r
36	380	25.2	861	6	AX042312 Sequence
37	380	25.1	834	1	AF368011 Cowdria r
38	379.5	25.1	825	1	AF368008 Cowdria r
39	378.5	25.1	834	1	AF368010 Cowdria r
40	378.5	25.0	863	1	AF355200 Cowdria r
41	378	25.0	873	1	AF028378 Cowdria r
42	378	25.0	1342	1	AF393388 Ehrlichia
43	376.5	24.9	2362	1	AF319940 Cowdria r
44	376.5	24.9	11329	1	AF082744 Ehrlichia
45	376.5	24.9	11329	1	AF082744 Ehrlichia

ALIGNMENTS

RESULT 1
AF230642 AF230642 14759 bp DNA linear BCT 01-JUN-2000
LOCUS Ehrlichia chaffeensis clpx (clpx) gene, complete cds; 28-kda outer
DEFINITION Membrane protein gene cluster, partial sequence and unknown genes.
ACCESSION AF230642
VERSION AF230642.1 GI:8132823
KEYWORDS Ehrlichia chaffeensis.


```

CDS
8214..8591
/note="Orf378"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAF73415.1"
/db_xref="GI:8132832"
/transl_table="MNCYLKGLPLNLOIFYGIFSRNFONNTLIIIPNDKCGFYTTL
WDPALHVTYTLGSEYVRNFEDILYENIQCKLLINYSVLNQHKNLTIVPIPN
ARFESNIRVNTSINKSESYEC"
8581..9432
/gene="p28-7"
/gene="p28-7"
/note="Component of 28-kDa outer membrane protein
multigene family"
/transl_table=11
/product="p28-7"
/protein_id="AAF73416.1"
/db_xref="GI:8132833"
/transl_table="MSAKKKLFILGSLVCLVSLPTKSLNLLNNNTKCTGLYVS
/QVAPTYSFNSFLKETYTDITKELLGLADIKSITDITTKKFNIPYTKFQDNVVS
FSAAGVYSQSDSRVEYSEYEDFVKPNQYVYSEAFRYIALARGIDNLOKQYPTNK
YVVIKNNGLSVASIIINGCYDFSLNNLKVSPIYICVGFQGDIIIEFFSAVSFKFAYQGRV
GISYPLFSNMIIFADGYVHKVIGNKNNLVQHVVSLSNHPKSTFAVATLNVEYFGE
FGLKFI"
9441..10268
/gene="p28-8"
9441..10268
/gene="p28-8"
/note="Component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-8"
/protein_id="AAF73417.1"
/db_xref="GI:8132834"
/transl_table="MSKKNFTTIGATLIHMLPNI SPETINNNTKLGLYISGOYK
PGTSHFSKFSKYETNDIOLIGLRHNAISTSLNINTDNIPKVNITFSFSGA
IGSDPTGARFELEGSEEDVDPGCLIKDYRYEALARNPSCGSPSTNNITVMRN
DGVSTVIFNGCQIDIFKDLQEVSPYCVGVGGDFIEFFDALHKLKAYQKGLGNYHL
STQASVFDYGYHKVIGNQNNLVQHVASTDFGPVYAVATLNIGYFGEIGRLTF"
10278..11135
/gene="p28-9"
10278..11135

Alignment Scores:
Pred. No.: 2,6e-126 Length: 14759
Score: 1510.00 Matches: 284
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-846-808-1 (1-284) x AF230642 (1-14759)

Qy 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIleLeuPheThrPro 20
Db 2197 ATGAGTAAGAGGATGACCAAGATTTCTTTGGGGTATGTTAATATTTATTTACACCA 2256

Qy 21 HisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValcIlyleGlnTyr 40
Db 2257 CATATTTCTTTAGCATCATGTTTAAATGACCATATTTCTATGATGTTGGTATTCAGTAC 2316

Qy 41 LysProAlaArgGlnHisLeuSerLysLeuIleLeuLysGloSerAlaAlaAsnThrVal 60
Db 2317 AAGCAGCCGACACACACTCTCAAAACTTCTTATATAAAGAAAGCGCAGCCCAACACTGTA 2376

Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyIleLysAspAsn 80
Db 2377 GAAGTTTTTGGTTAAAAAAGAACTTACTAATGATCTATTAACTGGGTATTAAAGATAAT 2436
```

```

Qy 81 ThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly 100
Db 2437 ACTAATTTTAAATATAATAACACCCATATTATGAAATAATATCGTTTAGTGTTCCTCGGT 2496

Qy 101 IlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr 120
Db 2497 ATATTGGATATTACTATAACAAAAATTTTCAGAAATAGAACTCTGAACCTTTCTTTACGAAACT 2556

Qy 121 PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
Db 2557 TTCCACATAAAGAACACACGATATAAAGAATTTGATTTGGAAGAACATTTTCTCTCTTGA 2616

Qy 141 LysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
Db 2617 AAGGAAATAAGCGGTGGTAGTAACAATCTCGCAATAATAAATATGTTACCCCTAATAAAT 2676

Qy 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys 180
Db 2677 AATGGTATTTCACCTTACCTCAGCTCGATCAACGTTTGTATGATGTTGACGAGCTTAAA 2736

Qy 181 HisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
Db 2737 CATATAATAAAGCTATTTCATGTTTAGATTTGGAGTAGACACAATTTGATTTTCTCACT 2796

Qy 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
Db 2797 AAGTATACTACAAAGCTTTTCATACCAAGGAAATTAGGAGCTAGTTATATCTGTTCTCCT 2856

Qy 221 GlnValSerValPheIleGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLys 240
Db 2857 CAAGTATCGGTTTTTATAGAGGTTTATTATCATGCTTTTATTTGGTAAAAAATTTGAGAAA 2916

Qy 241 IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis 260
Db 2917 ATTCTGTAAATATCCATGTGACTATCCATCCACCAACACCAACAGCAGCCACAT 2976

Qy 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlycylSerIleGlyIle 280
Db 2977 GTTCACACTACAGCATTAGCTATGTTAAGTATTGGATATTACGGTGGGAGTATTGGAATA 3036

Qy 281 LysPheIleLeu 284
Db 3037 AAGTTTATATTA 3048

RESULT 2
ECU72291 27190 bp DNA linear BCT 02-APR-2001
LOCUS ECU72291
DEFINITION Ehrlichia chaffeensis strain Arkansas major outer membrane protein
ACCESSION U72291 AF021338
VERSION U72291.2 GI:13511827
KEYWORDS
SOURCE Ehrlichia chaffeensis.
ORGANISM Ehrlichia chaffeensis.
REFERENCE
1 Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
1 (bases 14844 to 21136; 21479 to 22234)
Ohashi, N., Zhi, N., Zhang, Y., and Rikihisa, Y.
Immunodominant major outer membrane proteins of Ehrlichia
chaffeensis are encoded by a polymorphic multigene family
Infect. Immun. 66 (1), 132-139 (1998)
JOURNAL 98084465
MEDLINE 9423849
PUBMED 9423849
REFERENCE
2 (bases 1 to 27190)
Ohashi, N., Rikihisa, Y., and Unver, A.
Analysis of transcriptionally Active Gene Clusters of Major Outer
Membrane Protein Multigene Family in Ehrlichia canis and E.
chaffeensis
Infect. Immun. 69 (4), 2083-2091 (2001)
JOURNAL 21153566
MEDLINE 11254561
PUBMED 11254561
REFERENCE
3 (bases 1 to 27190)
```

AUTHORS TITLE JOURNAL	Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Direct Submission Submitted (24-SEP-1996) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	gene	/translation="MNNKSRGQVLLFAAYSIKFIYINSCTSKFRLSVKCFVYNPLKIFYKTLFKLGKCFVNYNYSKTSNHLRSLFKLFGIIRTNFKISYLNFTNPRVLIIYKILKKNKNSSINYLNNKYTINPTATSFINSVNIIGIYPGFVNYNI"
REFERENCE AUTHORS TITLE JOURNAL	4 (bases 1 to 27190) Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Direct Submission Submitted (27-AUG-1997) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	CDS	/gene="omp-10" 3915..4808 /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein omp-10" /protein_id="AAK28661.1" /db_xref="GI:13511832"
REFERENCE AUTHORS TITLE JOURNAL	5 (bases 1 to 27190) Ohashi,N., Rikihisa,Y. and Unver,A. Direct Submission Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	gene	/translation="MSYAKVFIILICILILLVPSLSFAIVNDFLKDNIHGFIYIGQYKPGVPRNRFVLVTNNIRELMSSDEECSTIPHMVQSAQGTLPPEALEELAKGLHGYYLFTLPNTYTKNLLGAGVIGYSTHERVEVEAEYERFNLTAAGYLHKNFYEPFALATMDTKPHQSAEDKYYIMKNKTITLSPFIINACYDFILKTRNVAPLYCLGVGGNFIDFLQVSKFAYQAKVGISYFVSPNIAPFFIDGSPFHGLNNOFSDLPVVDYSSGGFTISAKFNANFLTSSIGIRFIS"
REMARK COMMENT	Sequence update by submitter On Apr 2, 2001 this sequence version replaced gi:2853584 gi:2853273.	CDS	4825..5682 /gene="omp-1p" /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein omp-1p" /protein_id="AAK28662.1" /db_xref="GI:13511833"
FEATURES Source	1..27190 /organism="Ehrlichia chaffeensis" /strain="Arkansas" /db_xref="taxon:945" 315..956 /codon_start=1 /transl_table=11 /product="hypothetical transcriptional regulator"	gene	/translation="MOKLYISFILSGLLLPKYVFCMHQNNIDGSYVTIKYQLTTPHFKNFYKEDFDFTQEPDIGAKITANTKFTLKENFSPLHOTDSKYQNDLIGLGFNFKYKRIREFEGAYKNFTKRLARYKSKDGYKFAIPRKSEHGFLODNFTGYTAKNNGIISIINILNCSTKYKATPYICIGVGDFIEIFDVMRIKFAQCKVGYSPITSKLIIISQYHKVIGNKFELLVPYQPVVELKRLVTNKTSKDIDQDVTASLTMLEHFSSEIGUSFIF"
CDS	1136..2032 /gene="omp-1m" 1136..2032 /gene="omp-1m" /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein omp-1m" /protein_id="AAK28658.1" /db_xref="GI:13511829"	CDS	5695..6513 /gene="omp-1t" 5695..6513 /gene="omp-1t" /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein omp-1t" /protein_id="AAK28663.1" /db_xref="GI:13511834"
gene	2498..3352 /gene="omp-1n" 2498..3352 /gene="omp-1n" /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein omp-1n" /protein_id="AAK28659.1" /db_xref="GI:13511830"	gene	/translation="MYMYKKHYCYIVTVITLFLFLLPIESLSALIGNVEKDLKYSSTVSSQYKPSIFHFNFISOEHPKKSSEEFKKKANLNILKSNAYNQFQDNTTSFSGTIGYFSGLEAEAGCYOEFNVKNNSLIISNKKYHSRIHDEYAITTNKKLSIASIMVNTCYDISINNTSIVPYLCTGIGEDLVGLFNTTHFKLAYQCKVGMGYLINNNILLFSDIYYHKYMGNRKKNLYMQYVADPNISEETIPILAKLDIGYFGSGEIGIRFMEN"
CDS	3459..3905 /codon_start=1 /transl_table=11 /product="unknown function UN2" /protein_id="AAK28660.1" /db_xref="GI:13511831"	CDS	6529..7416 /gene="omp-1u" 6529..7416 /gene="omp-1u" /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein omp-1u" /protein_id="AAK28664.1" /db_xref="GI:13511835"
gene	7438..8277 /gene="omp-1v" 7438..8277 /gene="omp-1v" /note="omp-1 family member" /codon_start=1 /transl_table=11	gene	/translation="MTKFNFMVNILTFLLFLPPLKSFTTYANNNTITQKVGLYISGQYKPSIPHFKNFSVEENDKVVDLIGLTDVTYITEHLRDNTRKFNTHYIAKFKNNFINFSAIGYVSGQGRLEIESYGDVDVYKNYAVQDVRVYFALVREKNGSNFSPKPHETSQPSDSNPKSKSYTLLMKNNGYFVASVIINGCDYDFSNNTTISPYVICIGVGDIIEFEFVWHIRFACQSKVGISYPISPSITIFADAHYHKVINNNKFNHNVKYSTELKNSPTITISA TKALNIEFGGEGVMRFIF"
CDS		CDS	7438..8277 /gene="omp-1v" 7438..8277 /gene="omp-1v" /note="omp-1 family member" /codon_start=1 /transl_table=11

/product="major outer membrane protein omp-1v"
/protein_id="AAK28665.1"
/db_xref="GI:13511836"

/translation="MSKKKEITIGTVLASLLSFLSIESFSAINHNHTGNNSTGIYITG
QYRGVSHFSNFSKETSNDVDIQLGVKKSSSIDPNYSNFQGPYIVTFQDNRAFSF
GAIGTSPESURLESGYERKFDVKDPKDYSAKDAFFALARNITSTVPPDAQKTYVM
KNGLSVASIMINGCYDLSFNVLVSPYICAGIGEDFIEFFDTLHLKAYQCKLGISY
YFFPKINVAGGYHRVIGNFKNLNVNHVVTLDFFPKATSAVATLNVAYFGGAGVK
FTF"

CDS 8300..8890

/codon_start=1
/transl_table=11
/product="unknown function UN3"
/protein_id="AAK28666.1"

/db_xref="GI:13511837"
/translation="MIYKEKLTREGEYLAVLSFTLSTYIFLVNLNIRNSLAICVI
SLLRNFVSTKLIKDKCRDTRFSNMNLYLKGPLNLIQIFYGIFSRNFQNTLI
IPNSKCGFYTLWDNPNALHYTYLTGSEYRNFEDILYENIIOCKLLINYSVLNO
HNKNTLVIPNAREFSNEIRVRNISINKESSEYEC"
8880..9731

gene

CDS 8880..9731

/gene="omp-1w"
/note="omp-1 family member"

/codon_start=1
/transl_table=11
/product="major outer membrane protein omp-1w"
/protein_id="AAK28667.1"

Alignment Scores:

Pred. No.:	5,45e-126	Length:	27190
Score:	1510.00	Matches:	284
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-846-808-1 (1-284) x ECU72291 (1-27190)

Qy	1	MetSerLysArgSerAsnArgLysPheValLeuTvpValMetLeuLeuPheThrPro	20
Db	2498	ATGAGTAAAGAGAGTAAACAGAAGTTGTTTGGGGTATGTTAATATTATTACACCA	2557
Qy	21	HisLeSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValGlyIleGlnTyr	40
Db	2558	CATATTCTTTAGCATCAGTTTAAATGACCATAATCTATCTATGTTGGTATTCAGTAC	2617
Qy	41	LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaSerThrVal	60
Db	2618	AAGCCAGCCAGACACCTATCAAACTTCTTATAAAGAAAGCCAGCCACACTGTA	2677
Qy	61	GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyIleLysAspAsn	80
Db	2678	GAAGTTTGGGTTAAAAAGACCTACTAATATGATCTATTAACTGGTATTAAAGATAAT	2737
Qy	81	ThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly	100
Db	2738	ACTAATTTTAATATAAAATACACCATATTATGAAATAATACGTTTTCGCGT	2797
Qy	101	IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr	120
Db	2798	ATATTGGGATATTACTATAACAAAAATTCAGAATAGAACTCGAACTTCTTACGAAACT	2857
Qy	121	PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla	140
Db	2858	TTCCACATAAAGAACACGGATATAAAGAAITGATGTGGAAGCAATTTGCTCTTGA	2917
Qy	141	LysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn	160
Db	2918	AAGGAAATAAGCGGTGGTAGTAACAATCCTGCCAATAATAATATGTTACCTAATAAT	2977
Qy	161	AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys	180

Db	2978	AATGGTATTTCACCTACCTACGCTCTGATCAACGTTTGTATGATGATGTCGACGACTAAA	3037
Qy	181	HisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer	200
Db	3038	CATAATATAAATACGTTATTCATGTTTAGGATTTGGAGTAGACACAAATGATTTTCACGT	3097
Qy	201	LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyValAserTyrThrValSerPro	220
Db	3098	AAGTATACACAAAGTTTTCATACCAAGCAAAATAGGAGCTAGTTATATACTGTTCTCCT	3157
Qy	221	GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys	240
Db	3158	CAAGTATCGGTTTTATAGAAGGTTATATCATCTGTTTATTTGGTAAAAAATTTGAGAA	3217
Qy	241	IleProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHis	260
Db	3218	ATTCTGTAAATATCCATGTCATCTATCCATCACCACCAACCAACCAAGCAAGCCACAT	3277
Qy	261	ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIle	280
Db	3278	GTTCACTACACGATAGCTATGTTAAGTATTGGATATTACGGTGGAGTATTGGAATA	3337
Qy	281	LysPheIleLeu 284	
Db	3338	AAAGTTTATATTA 3349	
RESULT	3		
LOCUS	AF078553	28254 bp	DNA linear BCT 02-APR-2001
DEFINITION	Ehrlichia canis major outer membrane protein P30 multigene cluster 1, complete sequence.		
ACCESSION	AF078553 AF078554 AF078555 AH006958		
VERSION	AF078553.2 GI:13512584		
KEYWORDS	Ehrlichia canis.		
SOURCE	Ehrlichia canis.		
ORGANISM	Ehrlichia canis.		
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		
AUTHORS	1 (bases 1 to 28254)		
TITLE	Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y. Cloning and characterization of multigenes encoding the immunodominant 30-kilodalton major outer membrane proteins of Ehrlichia canis and application of the recombinant protein for serodiagnosis		
JOURNAL	J. Clin. Microbiol. 36 (9), 2671-2680 (1998)		
MEDLINE	98371112		
PUBMED	9705412		
REFERENCE	2 (bases 1 to 28254)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Analysis of Transcriptionally Active Gene Clusters of Major Outer Membrane Protein Multigene Family in Ehrlichia canis and E. Chaffeensis		
JOURNAL	Infect. Immun. 69 (4), 2083-2091 (2001)		
MEDLINE	21153566		
PUBMED	11254561		
REFERENCE	3 (bases 1 to 28254)		
AUTHORS	Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-1998) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA		
REFERENCE	4 (bases 1 to 28254)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA		
REMARK	Sequence update by submitter		
COMMENT	On Apr 2, 2001 this sequence version replaced gi:3790556 gi:3790555		
FEATURES	Location/Qualifiers		
source	1..28254		
	/organism="Ehrlichia canis"		
	/strain="Oklahoma"		
	/db_xref="taxon:944"		

CDS	<1. .615 /codon_start=1 /transl_table=11 /product="hypothetical transcriptional regulator" /protein_id="AAK28679.1" /db_xref="GI:13512585" /translation="PYANEIKAHDPDIEKNIQOQKFTGEGRLDWDVYASKNF KKEGASTKELKATVRHPDVECKEIKRORIMRGMSONLANKLIGITFOOVOKY EKGTRIVISRLYQLASVLNVEVRDMLKLQEDLKNISCDNPITPPHALNDNEKEFLP EPNDSKIDSKVLEMMVRATCTIKNEKVRNIYINLVKALSLDNKS" 843. .1733 /gene="p30-19" /note="p30-19" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-19" /protein_id="AAK28680.1" /db_xref="GI:13512586" /translation="MKLLYHLDNIMIKSAIGIVFSFIALFAPNAPSPVPIDFSNES EMAGFVAQYINIGFRPSPISAKYKTDEKSEKELTFLSKEETETIDUKKAGDFRKG YSYVARNYTGSGAIGYGGGLRVELEGSFRFDVDKOKYKNDPGRHVPALSKDSEI QNSSGSSNNKDYVVMKNEGFAISLMFNACYMDIIGRSSLPVNAICIGIQGIIIRFL GGTNIHTLAKLGLGLISPKTILFANGYVVKAKDNAPTNLVQYVVEISAAPKHD PIVYNADNAGCEVGLRFL" 2491. .3318 /gene="p30-18" 2491. .3318 /gene="p30-18" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-18" /protein_id="AAK28681.1" /db_xref="GI:13512587" /translation="MNNKLSLAYIALILFTSHVSSALVNDHNLVYFGIOYKPARHHL SNLLIKESQDVVEVLAKYDAIGSPDSTKEVNNFTIKYHPVDNNRLGFSVIFGYY YKNFRIESEIHEIQLNKEGHRVGPKEYPALKAPPSSQGYRHYVLINNGISTY SALINACYDILPAHNIITYSCIGFDIVIDFLSKYTRFSGQKGLGASPTISHRMSV FTEVYVHGLFGKFEQLPLNYNANTSPPOQPHVHTTASAILSIGYGVSGVGIKFL" 3425. .3871 /codon_start=1 /transl_table=11 /product="unknown function U2" /protein_id="AAK28682.1" /db_xref="GI:13512588" /translation="MNNKVQLNEVLLLLALPYSLHNSIDIEYNEQPIKKLSRLNKNTI KSQKHLNKKYILLKLTSSIGWFLKFSNDYSNQVIYLRTSLSKLVSIALFDFFK NIFLNYKVLTDLSLRNFSANYFYSKHHNITNCFVSRAYFYFLSCSY" 3882. .4832 /gene="p30-17" 3882. .4832 /gene="p30-17" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-17" /protein_id="AAK28683.1" /db_xref="GI:13512589" /translation="MNYAKVFILVCIIFLPPSLFSFATNNNYPVHEIGKSGHGFYIGVQ YKPGTPHFRFSIADOSTENLLAISHTKDYLFYSYEVGRGLFSLPQEQNLLHYATGG STLNTLKSNKFIPOGNYPTDNLGVGIVGYSINLRLEAEAFYEFKNKAPGY NYDTEYATAVYKGTQPVHYCHMKNTGIIILSSELVANTCYDTFLKIAKIAPIYLCI VGCGDIFDLQTRLRASQAKGLYALSPNLTFFVDGSGFHGMNNOFPGLLVDPYPT DISVSMGDNATYASEFTTMLAKLNMIFLAGSIGIRFIS" 4848. .5699 /gene="p30-16" 4848. .5699 /gene="p30-16" /note="p30 family member" /codon_start=1 /transl_table=11	gene
CDS	5716. .6510 /gene="p30-15" 5716. .6510 /gene="p30-15" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-15" /protein_id="AAK28685.1" /db_xref="GI:13512591" /translation="MVNLIINVIKYTIALAFLLLPRVSFSLICNIEKSIKLSVHIN SQKPSISQISNYLIQENNSKEKINILNSNTITYNMQLENSTTNFRFIIGYFFKR LRFVDSYEEFHDKDLSKANLSKYRMYNEDFONFTIADNKLUSITSAIVNICY DILINNTVPLHLCITAVGICSTGFFNDMRKLLYQRIKGLYLINSNMLFFNYYHK VMRNKLNLTQYSVDINAFDAITVLTANDDIGYFGEVGVRFIFN" 6525. .7394 /gene="p30-14" 6525. .7394 /gene="p30-14" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-14" /protein_id="AAK28686.1" /db_xref="GI:13512592" /translation="MLQRLNFINIILAFLLLLFPFQSFTLYIHDEITQNVGLYISQ YKPSIYGFKNFLIEENSHKTVELMGANDVHTVEYVKDNTKFTNTPYSKFRNSLIN LSGAIYSGOGPRLEIEGSEYFNQVASCNKPVKANRYIALVRDKPCGNIPQDHS HSNNKYTFIKNNGISITLSVMINGCYDIAESNVKISPYVCAGIGGDFITIFEMHIKF AYCGKGISILVSPSISIFANGHYHKVMDNVFNHLVKYIYIKLQDAPTITSARAKLRI YGFGEVGVRFVF" 7419. .8255 /gene="p30-13" 7419. .8255 /gene="p30-13" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-13" /protein_id="AAK28687.1" /db_xref="GI:13512593" /translation="MNNKSLILGTLISLLFSSLPKAFSVINHSIDSSNVNGLYFTG OYRPVSHFSGFTVRETNIAQOLVSLNTKNENHIIERTNFSGLYAKFODNRAAPS GAIGYSPEGLKFEIEISYKFGVKSTKNYQSTNAVIFALARQTTSSNPSDNKYVHK NSGLSVASVMINGSNYSFNLVSPYICAGIGEDFIEFFDTLYIKLAYOGKLGVNYS LSSRFNIFADMYHKVIGNQFKNLNVIHAVALDTPPKVTSAIATLNVAYPEGGEVGIHF IL" 8266. .8823 /codon_start=1 /transl_table=11 /product="unknown function U3" /protein_id="AAK28688.1" /db_xref="GI:13512594" /translation="MATYMIHKAKITPEMPSPSLTDYIILSLNNVDHNSFTICIT KLYKLILSSITKKLINVKDNTAVLNHLYTLNKSCKFSQTSQSFEMYPYTKIKOR ILNFSLIYIKINVTOKNSNESYKSYKSHNFNAYITTSFIHQHLLKYKYSTNOH NKSTKFSFNILAKNILLNNOENSYE" 8816. .9685 /gene="p30-12" 8816. .9685 /gene="p30-12" /note="p30 family member" /codon_start=1 /transl_table=11	gene
CDS	8266. .8823 /codon_start=1 /transl_table=11 /product="unknown function U3" /protein_id="AAK28688.1" /db_xref="GI:13512594" /translation="MATYMIHKAKITPEMPSPSLTDYIILSLNNVDHNSFTICIT KLYKLILSSITKKLINVKDNTAVLNHLYTLNKSCKFSQTSQSFEMYPYTKIKOR ILNFSLIYIKINVTOKNSNESYKSYKSHNFNAYITTSFIHQHLLKYKYSTNOH NKSTKFSFNILAKNILLNNOENSYE" 8816. .9685 /gene="p30-12" 8816. .9685 /gene="p30-12" /note="p30 family member" /codon_start=1 /transl_table=11	gene
CDS	8266. .8823 /codon_start=1 /transl_table=11 /product="unknown function U3" /protein_id="AAK28688.1" /db_xref="GI:13512594" /translation="MATYMIHKAKITPEMPSPSLTDYIILSLNNVDHNSFTICIT KLYKLILSSITKKLINVKDNTAVLNHLYTLNKSCKFSQTSQSFEMYPYTKIKOR ILNFSLIYIKINVTOKNSNESYKSYKSHNFNAYITTSFIHQHLLKYKYSTNOH NKSTKFSFNILAKNILLNNOENSYE" 8816. .9685 /gene="p30-12" 8816. .9685 /gene="p30-12" /note="p30 family member" /codon_start=1 /transl_table=11	gene

/product="major outer membrane protein p30-12"
/protein_id="AAK28689.1"
/db_xref="GI:13512595"

/translation="MNSKKTFSILGSILICLAACLPQSFSESSNVNTYTKHTGLYIS
GLYKPSVSHFSDQFSKEYNTNEALFGLKQDSSILRNKRTQYNNRNVPYTKAFQD
DFASPIAGYIANGPRIEIGSEEDVKNPGNYTIDAHRYTALAREKTSYLLSS
PKENRYVTKNGISIVIIINGCYDILSNDSPKSPYICTGGDFIEFFSAIRKFA
YQKIGISYSSLSNIIILTDGYHKVINSQFKNLNVHVNELTTPDKVTSATAFLNI
EYFGGEGLKLF"

Alignment Scores:		1.08e-73	Length:	28254
Pred. No.:	Score:	932.00	Matches:	180
Percent Similarity:		77.14%	Conservative:	36
Best Local Similarity:		64.29%	Mismatches:	58
Query Match:		61.72%	Indels:	6
DB:		1	Gaps:	2
US-09-846-808-1 (1-284) x AF078553 (1-28254)				
Qy	5	SerAsnArgLysPheValLeuTrpValMetLeuIleLeuPheThrProHisIleSerLeu	24	
Db	2494	ATAATAAATCTCCCTCTTATATGCTAATAATATATATATACATGATGTTCTTCA	2553	
Qy	25	AlaSerValLeuAsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArg	44	
Db	2554	GCATTGGTTTAAACGATCACAACTCTGTATATTTGGTATTCAATATAAACGAGCTAGG	2613	
Qy	45	GlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGly	64	
Db	2614	CATCATCTATCAATCTCTTATCAAGAAAGTAAGTCAGATGTTGTAGAGTGCCTGCA	2673	
Qy	65	LeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsn	84	
Db	2674	CTGAATATGATGCATAGTAGTCCATTCGACAGCAGTAAAGAGTAAATATTTTACC	2733	
Qy	85	IleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyr	104	
Db	2734	ATAAATAACAATCCACATATGATAATAATAGTTAGGCTTCTCTGTAATATTTGCTAT	2793	
Qy	105	TyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys	124	
Db	2794	TACTACAATAAATAATTTTAGTAGAATAACAGAAATTTCCACGAAATTTTCCAGCTAAA	2853	
Qy	125	AsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSer	144	
Db	2854	ATGAAGGACATAAAGAGCTTCGATTCGAGAAATATTTTGCCTAAG-----TTT	2904	
Qy	145	GlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSer	164	
Db	2905	GCTCCACCGCTCTACGCAAGGATATACACATGTTACTTTAATAACAATGATATTTCA	2964	
Qy	165	LeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIle	184	
Db	2965	ACCACCTTCAGCTTATTAATGCTGCTATGATGACTCATACCTGCACATATAATAA	3024	
Qy	185	ThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThr	204	
Db	3025	ACATATTATGTTAGGGTTTGAATAGATATAGATATTTCTAAGTAATAACACACTACA	3084	
Qy	205	LysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerVal	224	
Db	3085	AAATTTTCACATCAAGCAAGCTAGGAGCTAGTACCCTATTTCTCATAGATGTCAGTC	3144	
Qy	225	PheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsn	244	
Db	3145	TTTACAGAGGTTTATACCATGTTTATTCGCTAAATAATTTGACGAGCTTCTCTTAAAC	3204	
Qy	245	TyrProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThr	264	
Db	3205	TATAATGCT-----AATACATACACCAACCAACCAACCAACCAACCAACCAAC	3255	
Qy	265	AlaLeuAlaMetLeuSerIleGlyTyrTyrGlyLysSerIleGlyIleLysPheIleLeu	284	

Db	3256	GCATCGCTATATTAAGTATTGGCTATTATGTTGGAAGTGTTCGATAAAATTTATATTG	3315	
RESULT 4				
AF082744				
LOCUS		11329 bp	DNA	linear BCT 18-SEP-2000
DEFINITION		Ehrlichia canis p28 multigene locus, partial sequence.		
ACCESSION		AF082744	AF168788	AF168789
VERSION		AF082744.2	GI:10181081	
KEYWORDS		Ehrlichia canis.		
SOURCE		Ehrlichia canis		
ORGANISM		Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;		
		Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		
REFERENCE		1 (bases 1 to 11329)		
AUTHORS		McBride,J.W., Yu,X.J. and Walker,D.H.		
TITLE		Molecular cloning of the gene for a conserved major immunoreactive		
		28-kilodalton protein of Ehrlichia canis: a potential		
		serodiagnostic antigen		
		Clin. Diagn. Lab. Immunol. 6 (3), 392-399 (1999)		
JOURNAL		99242757		
MEDLINE		10225842		
PUBMED				
REFERENCE		2 (bases 1 to 11329)		
AUTHORS		McBride,J.W., Yu,X.J. and Walker,D.H.		
TITLE		A conserved, transcriptionally active p28 multigene locus of		
		ehrlichia canis		
JOURNAL		Gene 254 (1-2), 245-252 (2000)		
MEDLINE		20432107		
AUTHORS		3 (bases 1 to 11329)		
TITLE		McBride,J.W., Yu,X.J. and Walker,D.H.		
JOURNAL		Submitted (07-AUG-1998) Pathology, University of Texas Medical		
		Branch, 301 University Blvd., Galveston, TX 77555-0609, USA		
REFERENCE		4 (bases 1 to 11329)		
AUTHORS		McBride,J.W., Yu,X.J. and Walker,D.H.		
TITLE		Submitted (04-AUG-2000) Pathology, University of Texas Medical		
JOURNAL		Branch, 301 University Blvd., Galveston, TX 77555-0609, USA		
REMARK		Sequence update by submitter		
COMMENT		On Sep 18, 2000 this sequence version replaced gi:3769522.		
FEATURES		Location/Qualifiers		
		1..11329		
		/organism="Ehrlichia canis"		
		/strain="Jake"		
		/db_xref="taxon:944"		
		<1..146		
		/gene="p28a"		
		<1..146		
		/note="outer membrane protein"		
		/codon_start=3		
		/transl_table=11		
		/product="p28"		
		/protein_id="AAAG14355.1"		
		/db_xref="GI:10181083"		
		/translation="SIISHSVKNLPVQVNTLEEYPRVTSIAITLDIGYLGGEIGIR		
		EIF"		
		170..1051		
		/gene="p28-1"		
		170..1051		
		/gene="p28-1"		
		/note="outer membrane protein"		
		/codon_start=1		
		/transl_table=11		
		/product="p28-1"		
		/protein_id="AAAG14356.1"		
		/db_xref="GI:10181084"		
		/translation="MNNKLKFTIINTLVLCCLSLPNISSKAINNNKAYGLYISGO		
		YKPSVSVSNFSVKETNWTNKLIALKDDVDSIEKTQDASVGSINPSNPTIPYTAFFQ		
		DNSVFNQITGTFAGRTVEEGSEEDVKNPGYTLSDRYRTFALAREMKMSFT		
		PKKVSNTIFHTVMRDGLSIISVINVCYDFSLNLSISPYICGAGVDAIEFFDL		
		HIKFAVQSKGLIAYSLPSILSFASLYHKVMGNQFKNLNVQHVLAELASIPKITSVA		
		TLNIGYFGGEIGEARLTF"		

Db	269	AAAAAGTACTACCGATTATATATACAGTGGACAATAATAAACCCAGTGTCTGTTTTCAGT	328
Qy	49	LysLeuLeuIleLysGluSerAlaAlaAasnThrValGluValPheGlyLeuLysLysAsp	68
Db	329	AATTTTTCAGTTAAAGAACCAATGTCATAACTAAACACCTTATAGCTTTAAAAAAGAT	388
Qy	69	Leu-----LeuAasnAspLeuLeuThrGlyIleLysAspAsnThrAsnPhe	83
Db	389	GTTGACTCTATTGAAACCAAGACTGATGCCAGTGTAGTATTAGTAACCCATCAAAATTT	448
Qy	84	AsnIleLysTyrAsnProTyrTyrGluAasnAsnArgLeuGlyPheSerGlyIlePheGly	103
Db	449	ACTATCCCTTACAGCTGATTTTCAAGATAATCTGCTAATTTCAATGGAATCTGTT	508
Qy	104	TyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIle	123
Db	509	TACACCTTCTGAAGGTACAAAGTGAATAGAGGTCCTTATGAGGAATTTGATGTT	568
Qy	124	LysAsn---AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGlu	142
Db	569	AAAAACCTGGAGGTATACACTAAGTGATGCTATCGCTATTTTGCATTAGCACCTGAA	628
Qy	143	IleSerGlyClySerAsnAsnPro-----AlaAsnAsnLysTyrValThrLeu	158
Db	629	ATGAAAGGTAAATAGTTTACACCTAAGAAAGTTTCTAATAGTATTTTTCACACTGA	688
Qy	159	Ile---AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAsp	177
Db	689	ATGAGAAATGATGATTAUCTATAATATCTGTTATAGTAATGCTTGTACGATTTCTCT	748
Qy	178	GlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp	197
Db	749	TTGAACAATTTGCAATATCGCCTTACATATGTTGGAGGACGAGGTAGATGCTATAGAA	808
Qy	198	PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr	217
Db	809	TTCTTCGATGTAATACACATTAAGTTTGCATATCAAGCAAGCTAGTATTGCTTATCT	868
Qy	218	ValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLys	237
Db	869	CTACCATCAACATTAAGTCTCTTGTAGTTTATATACCTAAAGTAATGGCAATCA	928
Qy	238	PheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSer	257
Db	929	TTTAAAAATTTAAATGTCACACATGTTGCTGAACCTGCA-----AGT	970
Qy	258	LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySer	277
Db	971	ATACCTAAAAATT---ACATCCGAGTTGCTACACTTAATATGTTTATTGAGGTGAA	1027
Qy	278	IleGlyIleLys 281	
Db	1028	ATTGGTGCAAGA 1039	
RESULT 5			
AF324792			
LOCUS	AF324792	6913 bp	DNA linear BCT 11-APR-2001
DEFINITION	Ehrlichia canis phosphoribosylaminoimidazole carboxylase (purk) gene, complete cds; major outer membrane protein gene cluster 2, complete sequence; and u6 gene, partial cds.		
ACCESSION	AF324792		
VERSION	AF324792.1	GI:13591681	
KEYWORDS	Ehrlichia canis.		
SOURCE	Ehrlichia canis.		
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		
REFERENCE	1 (bases 1 to 6913)		
AUTHORS	Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.		
TITLE	Cloning and characterization of multigenes encoding the immunodominant 30-kilodalton major outer membrane proteins of Ehrlichia canis and application of the recombinant protein for serodiagnosis		

JOURNAL	J. Clin. Microbiol.	36 (9),	2671-2680 (1998)
MEDLINE	98371112		
PUBMED	9705412		
REFERENCE	2 (bases 1 to 6913)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Analysis of Transcriptionally Active Gene Clusters of Major Outer Membrane Protein Multigene Family in Ehrlichia canis and E. chaffeensis		
JOURNAL	Infect. Immun.	69 (4),	2083-2091 (2001)
MEDLINE	21153566		
PUBMED	11254561		
REFERENCE	3 (bases 1 to 6913)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-NOV-2000)	Veterinary Biosciences, Ohio State University, 1925 Coffey Rd., Columbus, OH 43210, USA	
FEATURES	Location/Qualifiers		
source	1..6913		
	/organism="Ehrlichia canis"		
	/strain="Oklahoma"		
	/db_xref="taxon:944"		
gene	329..1420		
	/gene="purk"		
CDS	329..1420		
	/gene="purk"		
	/codon_start=1		
	/transl_table=11		
	/product="phosphoribosylaminoimidazole carboxylase"		
	/protein_id="AAK31311.1"		
	/db_xref="GI:13591682"		
	/translation="MLHNSYVSPGSTIGIGGQLGKMISIAATNLGYKTHLLTDNQD SPFINVTNVTILDYNNKELLEFPASSVDIATLEFENIPSTPIDILSOKTKVYPGKT AHIHSONRIRKQYIRNLGIKTANFRIDYNDLVHNLVLELGYPALLKRTTELGYDCKG QYLITQCDLNTLSLWQOEYILEKFVKIHKELISIIISKSINGSIEFFPPVAENHHTN GLATSSVPASIEEINTQAKNIALQAESNLVGLIAVEFFITDQELIVNEIADRP HNSGHWSLDACNISQEOULTRAICGLPLKPVLLFPCTMDNLIGDDIINYQYKTKIN ESIIYVGTGKASKNRKMGHINRLNLKHEI"		
gene	1756..2598		
	/gene="p30-10"		
CDS	1756..2598		
	/gene="p30-10"		
	/note="P30 family"		
	/codon_start=1		
	/transl_table=11		
	/product="major outer membrane protein p30-10"		
	/protein_id="AAK31312.1"		
	/db_xref="GI:13591683"		
	/translation="MNYKKILVRSALISLSMILPYQSFPADVPGSRNTDNKGEFYISAK YNPSISHFRKFAEETPINGTNSLTKKVKLGKDGDTTKDDFTRVAPGIDFQNNLIS GFSGSIGYSGMDGPRIEAAAYQOFNPKNTDNDONGEYIKHFALSRKDMEDQOQVY LKNDGITFMSLVNTCYDITAEQSVFPVACAGIGADLITIFKDLNLKPAYOCKIGIS YPITPEVSAPIGGYHGVIGNKFEKIPVITPVVLNDAPQTTASVTLDVGYFGGEIGM RFTF"		
gene	2926..3756		
	/gene="p30-4"		
CDS	2926..3756		
	/gene="p30-4"		
	/note="P30 family"		
	/codon_start=1		
	/transl_table=11		
	/product="major outer membrane protein p30-4"		
	/protein_id="AAK31313.1"		
	/db_xref="GI:13591684"		
	/translation="MNCCKILITTLVSLTILLPGISFSKPTHENTTGNFYILIGKYV PSISHGFNSAKEERTTIGIFGLKESWTGGIILDKREHAFNIPNTSFYENNPFLGK AGVIGYSIGSPRIEVEYETDVQNPQDKFNDAHKYCALNSDSKTKMKSGKFLFK NEGLSDISLMNVHYDIINKRMPFSYICAGIGTDLIFMFDATINHKAAYQGLGFNPV ISPEANISMGVHFHKVTNNEFRVPVLLTAGGLADPNLFAIVKLISICHGFLGFCYRVSE "		
gene	3967..4830		
	/gene="p30a"		
CDS	3967..4830		
	/gene="p30a"		

```
/note="P30 family"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30a"
/protein_id="AAK31314.1"
/db_xref="GI:13591685"
/transl_table="MKYKKFTVTALVLLTSFTHFIPFYSPARASTIHNFYISGKYP
TASHFGIFSAKEQSFVKVLGDORLSHNIINNDDAKSLKQVNYFKYKNNPFLGF
AGATGYSIGNSRLEYSHEIDFTKPNPNVLDNSHKYCALSHGSHICSPGNSGDWYT
AKYDKVLLKNEGLDVSPEMLNACDYLTTEKMPSPSYICAGICTDLDLSMRETTQNKIS
YQKLGUNYINRSVFAAGHFHKVIGNEFKGIPTLPLDGSNIKVQQAFAVTLDVCH
FGLIEGSRFF"
gene complement(5309..>6913)
CDS /gene="u6"
complement(5309..>6913)
/gene="u6"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK31315.1"
/db_xref="GI:13591686"
/transl_table="KLIVTQDKRVSDEIVDFHGISIVRPKINFNDHQFYLYONCKDK
VLKIDEXCDQNAVQIQRXQDDKVTYKCIDGINFTFGYVLKSHNVYDKHTWLPKPNK
MYRVLDKQIYQCADYEDLNKNQDFLOSLINDNINFTFGYVLKSHNVYDKHTWLPKPNK
DNPENLNKSFYLYQNTKLIKADNKSFTLPEVFTQFIKSDLSNVEYFNKLNQKLYD
SADPMIKSYRFDTHLVFNVDNKMESLDPKIQLEQPLDYHSMGMCIDSFESTTVY
PEKMTESDKNFDLALPVLOEDNKSESVLLHVPKLSGIYVPGKDFIKVEMLGGESG
KIDIKSWFKGEGOPGEVGMFKTKSKDKYFVKYPRKMKNSKSKSENKHLKDNSDV
GMDAVIEFCKDKRHDGVECVQLTANGGKLOHDIRKNVKSNNLVYDKRMLYRYVY
NSHVSRSSENVPHDKRIRFIYQNFITEALWEEIEKDECKSGKLLAENNSKYFGAGGC
VDVSTGSTERGASGMVKITCEDWYHE"
BASE COUNT 2520 a 1149 c 888 g 2356 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 1.4e-28 Length: 6913
Score: 425.00 Matches: 104
Percent Similarity: 51.92% Conservative: 45
Best Local Similarity: 36.24% Mismatches: 100
Query Match: 28.15% Indels: 38
DB: 1 Gaps: 11
```

US-09-846-808-1 (1-284) x AF324792 (1-6913)

```
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAla-----SerValLeu 28
::: |||::: |||::: ||| ||| |||
Db 1792 ATCTCATTAATGTCATCTTACCATATCAGCTTTTGCAGATCCTGTAGGTCAGAAGCT 1851
Qy 29 AsnAspHis----AsnSerMetTyrValGlyIleGlnTyrLysProAlaAargLnhIstLeu 47
|||::: |||::: |||::: |||::: |||
Db 1852 AATGATACAAAGAAGAGGCTTCTACATAGTCGAAAGTACAATCCAAAGTATATCACACTTT 1911
Qy 48 SerLysLeuLeuIleLysGluSerAlaAlaAsn-----ThrValGluVal 62
||| ||| |||::: ||| ||| |||
Db 1912 AGAAATTTCTCTGCTGAAGAAGCTCTTATTAATGAACAATTTCTCTCATAAAGT 1971
Qy 63 PheGlyLeuLysLysAsp-----LeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
||| ||| ||| ||| ||| |||
Db 1972 TTCGACTTAAGAAGATGGTGATATACAAAAAGACGATTTTACAGAGTAGCTCCA 2031
Qy 80 AsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
||| ||| ||| ||| ||| |||
Db 2032 GCGATTGATTTT-----CAAAATACTTAATATCAGGATTT 2067
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
||| ||| ||| ||| ||| |||
Db 2068 TCAGGAAGTATGTGTTACTCTATGAC---GGACCAAGAATAAGAACTTGAAGTCGATAT 2124
Qy 119 GluThrPheHisIleLys-----AsnAsnGlyTyrLysArgIleAspCysGluLys 135
||| ||| ||| ||| ||| |||
Db 2125 CAACAATTAATCAAAAAACCCGATACCAATGATGATCTGATATGTTGGTGAATACTATAA 2184
Qy 136 HisPheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAlaAsnLysTyr 155
```

```
Db 2185 CATTTGTCATTATCTCGTAAA-----GATGCAATGGAAGATCAGCAATAT 2229
Qy 156 ValThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystTyrAsp 175
||| ||| |||::: |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2230 GTAGTACTAAAAATGACGGCATACTTTTATGTCATTGCTGGTTAACTTGCATATGAC 2289
Qy 176 ValAspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThr 195
::: |||::: |||::: |||::: |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2290 ATTACAGCTGAAGGAGTAGTCTTTCGTACCATGATGTCAGGATATAGGACGACATCTT 2349
Qy 196 IleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSer 215
||| ||| |||::: |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2350 ATCACTATTTTAAAGACCTCAATCTAAAAATTGCTTACCAAGGAAAAATAGGTATTAGT 2409
Qy 216 TyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGly 235
||| ||| |||::: |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2410 TACCTATCATCACACCAAGAGTCTCTGCATTTATTTGTGTGATACCTACCGCGCTTATGCT 2469
Qy 236 LysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProPro 255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2470 AATAAATTGAGAAGATACCTGTA-----ATAACTCTGTAGTATTA 2511
Qy 256 AsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGly 275
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2512 AATGATGCTCCTCAAACC---ACATCTGCTTCAGTAACCTTTCAGCTTGGATACTTTGCG 2568
Qy 276 GlySerIleGlyIleLysPhe 282
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2569 GGAGAAATTGGAATGAGGTTT 2589

RESULT 6
AF125274 3535 bp DNA linear BCT 20-APR-1999
LOCUS Cowdria ruminantium isolate Welgevonden major antigenic protein 1
DEFINITION Like protein and major antigenic protein 1 (MAP1) genes, complete
cds.
ACCESSION AF125274
VERSION AF125274.1 GI:4589094
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Cowdria ruminantium
REFERENCE 1 (bases 1 to 3535)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE The map1 Gene of Cowdria ruminantium Is a Member of a Multigene
Family Containing Both Conserved and Variable Genes
JOURNAL Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)
REFERENCE 2 (bases 1 to 3535)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) College of Veterinary Medicine Department
of Pathobiology, University of Florida, P.O. Box 110880,
Gainesville, FL 32610, USA
FEATURES
source Location/Qualifiers
1..3535
/organism="Cowdria ruminantium"
/isolate="Welgevonden"
/db_xref="taxon:779"
987..1835
/note="ORF2"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1 like protein"
/protein_id="AAD26343.1"
/db_xref="GI:4589095"
/translation="MNYKKILVRSALISLMSFLPYQSFAPVSSNNIGNENAKEGFYI
SAKNVPSIHPHFKEPSABETPVYKDSPTKVFGLKKGSIITKYSDFTDTISFGQNN
FISGFSIGYIMDGPVETEAAYOKENPKPANETDTSYKYHGLSRAEMADKKY
VLTNNGVTSSLMFNACYDITAGVFPFIPYACAGIGADLISIFDDINLKPAYQCKIG
ISYPTTPEISAFIGGYHGVIGNKYNKVPVKLPVLTDAPOSTASVTLADAYFGGEL
GVRETF"
```

[illegible]

```
Alignment Scores:
Pred. No.: 7.67e-29 Length: 3538
Score: 424.00 Matches: 98
Percent Similarity: 50.87% Conservative: 49
Best Local Similarity: 33.91% Mismatches: 102
Query Match: 28.08% Indels: 40
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF125277 (1-3538)

Qy 13 ValMetLeuLeuPheThrProHisLeuAlaSerValLeuAsnAspHisAsn 32
Db 107 ATCTCGTTAATGTCATTTTACCATATCAGTCTTTGCAGAACCTGTGAAGTCAATAAC 1086
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
Db 1087 ATTCGAATGAATGCTAAGAGGGTTCTACATAGTGAAGTGAAGTCAACCCAGCAT 1146
Qy 45 GlnHisLeuSerLysLeuLeuIleLysGluSerAla-----AlaAsnThr 59
Db 1147 CCACATTCACAAAATTTCTGCTGAGGAACTCTCTGTATACGGTAAAGACTCTCCAAC 1206
Qy 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
Db 1207 AAAAGGTTATTTGGGTTAAAAAGGAG-----GGTTCTATAACAAAATACAGTGAT 1257
Qy 80 AsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
Db 1258 TTCACTAGACAGATATATCGTTGAGGCG-----CAAAATATTTTATCTCAGGTTTC 1311
Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 1312 TCAGGAACATAGTTATATCATGAT---GGACCAAGAGTAGATGAAGTGCATAC 1368
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 1369 CAAAATTCACACCCAAAATAATCCAGCTAATGAACTAGTACTAGTATTAATAAACAC 1428
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 1429 TATGATTATCTCGT-----CGAAGAGCTATGGCAGATGAAAATATATGTT 1473
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuAsnValCysTyrAspVal 176
Db 1474 GTACTTACAAATAATGGAGTAACCTTTTCATCATTAATGTTCAATGCTGCTATGACATT 1533
Qy 177 AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
Db 1534 ACAGCAGAGAGGAGTACCTTTTCATTCATATGATGCTGCTGATTTGGTGCAGATCTTATA 1593
Qy 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 1594 TCTATATTGATGATATAAATTTAAATTTTCTTACCAAGGTAATTTGGTATTAGTTAT 1653
Qy 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
Db 1654 CCAATTACTCTGAAATTTCTGCATTTATTCGCGGATATTAACATGGAGTAATAGGTAAC 1713
Qy 237 LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr 253
Db 1714 AAGTATACAAAGTACCTGTGAAGCTTCTGTAACATTTTAACAGATGCTCTCTCAA----- 1767
Qy 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
Db 1768 -----AGCACTCCGCTTCAGTAACCTCTTGACGCTGGATAT 1803
Qy 274 TyrGlyGlySerIleGlyIleLysPhe 282
Db 1804 TTTGGTGGTGAACCTGGAGTAAGGTTT 1830

RESULT 8
AF125279
```

```
LOCUS AF125279 3541 bp DNA linear BCT 20-APR-1999
DEFINITION Cowdria ruminantium isolate Antigua major antigenic protein 1 like
protein and major antigenic protein 1 (MAP1) genes, complete cds.
ACCESSION AF125279
VERSION AF125279.1 GI:4589109
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE 1 (bases 1 to 3541)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE The map1 Gene of Cowdria ruminantium Is a Member of a Multigene
Family Containing Both Conserved and Variable Genes
JOURNAL Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)
REFERENCE 2 (bases 1 to 3541)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) College of Veterinary Medicine Department
of Pathobiology, University of Florida, P.O. Box 110880,
Gainesville, FL 32610, USA
FEATURES
source
1..3541
/organism="Cowdria ruminantium"
/isolate="Antigua"
/db_xref="taxon:779"
989..1837
/note="ORE2"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1 like protein"
/protein_id="AAD26353.1"
/db_xref="GI:4589110"
/translation="MNVKKILVRSALISLMSFLPYOSFAEPVSSNNIGNNAKEGFYI
SAKNPSIPHRKFSAEETPVYKDSPTKKYFKLKEGSIKYSDFTRIDISFGQNN
VIFGSGSIGYIMDGPVEIEAAYKFNPNPNETDTSYKYHYSRGRAMADKKY
VFLNNGVTSSLMFNACYDITAEGVPEIPYACAGIGADLISIFDDINLKFAYQKIG
ISYPTITPEISAFIGGYHVGIGNKYNKVPKLVLTLDPOSTSASVTLDAFYFGEL
GVRTF"
gene
2210..3055
/gene="MAP1"
2210..3055
/gene="MAP1"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1"
/protein_id="AAD26354.1"
/db_xref="GI:4589111"
/translation="MNCKKIPITSLISLVSFLPGVSFSDVIOEDSPVGSVYISAKY
MPTASHFGKMSIKEDSRDTKVFGKKDWGVKTSSTNTIFTEKDYSPKYNNPFLGF
AGAIGYSMNGPRIEFETSYETFDVKNPGNKNDAHMYCALDTSSTGGAASVSMV
KNEMLTISLMLNACYDMLDMPVSPVCAGICTDLVSLVNSTNPKLSYOGKLGISY
SINPEASIFIGGHRFVIGNFKDITTSKIFNTSNTGGATPGFASAILDVCHFGLIEIG
GRFVF"
BASE COUNT 1282 a 451 c 479 g 1329 t
ORIGIN
Alignment Scores:
Pred. No.: 7.68e-29 Length: 3541
Score: 424.00 Matches: 98
Percent Similarity: 50.87% Conservative: 49
Best Local Similarity: 33.91% Mismatches: 102
Query Match: 28.08% Indels: 40
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF125279 (1-3541)

Qy 13 ValMetLeuLeuPheThrProHisLeuAlaSerValLeuAsnAspHisAsn 32
Db 1075 ATCTCATTAATGTCATTTTACCATATCAGTCTTTGCAGAACCTGTGAAGTCAATAAC 1084
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
Db 1075 ATCTCATTAATGTCATTTTACCATATCAGTCTTTGCAGAACCTGTGAAGTCAATAAC 1084
```

```
Db 1085 ATTGGAATGAAATGCTAAAGAAGGTTCTACATAAGTGCAAAATACAAACCAAGCATA 1144
Qy 45 GlnHisLeuSerLysLeuLeuLeuLysGluSerAla-----AlaAsnThr 59
Db 1145 CCACACTTCAGAAAATTTCTGCTGAGGAACCTCTGTATACGGTAAAGACTCTCCAAC 1204
Qy 60 ValGluValPheGlyLeuLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
Db 1205 AAAAGGTATTGGTTAAAAAGGAG-----GGTCTATACAAATACAGTGTAT 1255
Qy 80 AsnThrAsnPheAsnLeuLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
Db 1256 TTCCTAGACACATATATCGTTGAGGCG-----CAAAATATTTATCTCAGGTTTC 1309
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 1310 TCAGGAAGCATAGGTATATCATGGAT---GGACCAAGAGTAGAGATTGAAGCTGCATAC 1366
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 1367 CAAAAATCAACCAAAATCCAGCTAATGAACTGATAGTACTAGTACTATAAACAC 1426
Qy 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 1427 TATGGATTATCTCGT-----GCAGAAAGCTATGCGAGATAAAAAATATGTT 1471
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 1472 GTACTTACAATAATGAGTAACCTTTTCATCATTAATGTTCAATGCTGCTATGACATT 1531
Qy 177 AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
Db 1532 ACAGCAGAGGAGTACCTTTTCATTCATATGCTGCTGCTGTTGGTGCAGATCTATA 1591
Qy 197 AspPheLeuSerLysTyrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 1592 TCTATATTTGATGATATAAATTTAAATTCGCTTACCAAGTAAATTTGTTATAGTTAT 1651
Qy 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
Db 1652 CCAATTACTCTGAAATTTCTGCATTATTTGGCGGATATTACCATGAGTAATAGGTAAC 1711
Qy 237 LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr 253
Db 1712 AAGTATAACAAAGTACCTGTAAGCTCTCTGTAACTTTAACAGATGCTCCTCAA----- 1765
Qy 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
Db 1766 -----AGCACTTCGCTTACGTAACCTCTTGACGCTGGATAT 1801
Qy 274 TyrGlyGlySerIleGlyIleLysPhe 282
Db 1802 TTGGTGGTGAACCTTGAGTAAGGTTT 1828

RESULT 9
AF125278 3572 bp DNA linear BCT 20-APR-1999
LOCUS Cowdria ruminantium isolate Umbanein major antigenic protein 1 like
DEFINITION protein and major antigenic protein 1 (MAP1) genes, complete cds.
ACCESSION AF125278
VERSION AF125278.1 GI:4589106
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE 1 (bases 1 to 3572)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE The map1 Gene of Cowdria ruminantium Is a Member of a Multigene
Family Containing Both Conserved and Variable Genes
JOURNAL Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)
REFERENCE 2 (bases 1 to 3572)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
```

```

TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) College of Veterinary Medicine Department
of Pathobiology, University of Florida, P.O. Box 110880,
Gainesville, FL 32610, USA
FEATURES
source
1..3572
/organism="Cowdria ruminantium"
/isolate="Umbanein"
/db_xref="taxon:779"
989..1837
/note="ORF2"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1 like protein"
/protein_id="AAD26351.1"
/db_xref="GI:4589107"
/translation="MNYKILVRSALISLMSFLPYOSFAEPVSSNNIGNENAKEGYFI
SAKNPSIPHERKFSAEETPVYKGDSPKTKKVKFKKGGSTIKYSIDTTRTDISEGQNN
FISGSGSIGYIMDPRVEIARAYQKFNKPANETDTSYKHYGLISRAEMADKKY
VLTNGVTFSLLMFNACYDITAEGVFPFIPYACAGIGADLISIFDDINLFAYQGGKIG
ISYPTPEISAFIGYHYGVIGNKYNKVPKLPVTLTDPQSTASVTLDAFYEGGEL
GVRFTE"
gene 2210..3082
/genes="MAP1"
CDS 2210..3082
/genes="MAP1"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1"
/protein_id="AAD26352.1"
/db_xref="GI:4589108"
/translation="MNCKIFITSLISLVFLPVGSFSDVIOEDSNPAGSVVISARY
PTASHFEGKMSIKEDSKNTQTVFGLKDMGVKTPSSDGSNSIITEKDYSFKYENN
PFLGPAAGATGYSMNGPRIEESYEFYDVKNGKNDAMHMCALDITPGSTQGNAT
LNSVMVKNELTDIALMLNACYDITLEGMPVSPYVCAGIGTOLVSVINATNPKLSIQ
GKLGISYINPEASIFIGGFHVRVIGNEFKDITTSKIFTSTGKLATAASPGFASATLD
VCHFGEIGGRFVF"
BASE COUNT 1303 a 472 c 476 g 1321 t
ORIGIN
Alignment Scores: 7.76e-29 Length: 3572
Pred. No.: 424.00 Matches: 98
Score: 50.87% Conservative: 49
Percent Similarity: 33.91% Mismatches: 102
Best Local Similarity: 28.08% Indels: 40
Query Match: 1 Gaps: 10
DB:
US-09-846-808-1 (1-284) x AF125278 (1-3572)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 1025 ATCTCATTAATGTCATTTTACCATATCAGTCTTTTGCAGAACCTGTAAGTTCAAATTAAC 1084
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
Db 1085 ATTGGAATGAAATGCTAAAGAAGGTTCTTACATAAGTGCAAAATACAAACCAAGCATA 1144
Qy 45 GlnHisLeuSerLysLeuLeuLysGluSerAla-----AlaAsnThr 59
Db 1145 CCACACTTCAGAAAATTTTCTGCTGAGGAACCTCTGTATACGGTAAAGACTCTCCAAC 1204
Qy 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
Db 1205 AAAAGGTATTGGTTAAAAAGGAG-----GGTCTATACAAATACAGTGTAT 1255
Qy 80 AsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
Db 1256 TTCCTAGACACACATATCGTTTGGAGGCG-----CAAAATATTTATCTCAGGTTTC 1309
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 1310 TCAGGAAGCATAGGTTTATATCATGGAT---GCACCAAGAGTAGAGATTGAAGCTGCATAC 1366
```

QY	119	GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis	136
Db	1367	CAAAATTCACACCCAAAATAATCCAGCTAATGAACATGACATAGTGAATTAATAACAC	1426
QY	137	PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal	156
Db	1427	TATGGATATCTCGT-----CCAGAAGCTATGGCAGATAAAAAATATGTT	1471
QY	157	ThrLeuIleAsnAsnAnglyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal	176
Db	1472	GTACTTACAAATAATGGAGTAACATTTTCATCAITTAATGTTTAATGCTGCTATGACAT	1531
QY	177	AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIle	196
Db	1532	ACAGCAGAAGGAGTACCTTTCATTCATATCATGTGCTGCTATGTTGGTCAGATCTTATA	1591
QY	197	AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr	216
Db	1592	TCTATATTGGATATATAAATTTAAATTTGCTTACCAAGGTAAATTTGGTATTAGTTAT	1651
QY	217	ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys	236
Db	1652	CCAACTACTCCTGAATTTCTGCATTTATTGGCGGATATTACCATGGAGTAATAGGTAAC	1711
QY	237	LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr	253
Db	1712	AAGTATAACAAGATACCTGTGAAGCTTCTGTAACATTTAACAGATGCTCCTCAA-----	1765
QY	254	ProIleAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr	273
Db	1766	-----AGCATTCCTCGCTTCAGTAACCTCTTGACGCTGGATAT	1801
QY	274	TyrGlyGlySerIleGlyLysPhe	282
Db	1802	TTTGGTGGTGNACTGGAGTAAGTTT	1828
RESULT	10		
AF125276			
LOCUS		3507 bp DNA linear	BCT 20-APR-1999
DEFINITION		Cowdria ruminantium isolate Highway major antigenic protein 1 like	
ACCESSION		protein and major antigenic protein 1 (MAP1) genes, complete cds.	
VERSION			
KEYWORDS			
SOURCE		AF125276.1 GI:4589100	
ORGANISM		heartwater rickettsia.	
REFERENCE		Cowdria ruminantium	
AUTHORS		Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;	
TITLE		Rickettsiaceae; Ehrlichiae; Cowdria.	
JOURNAL		1 (bases 1 to 3507)	
REFERENCE		Sulsona,C.R., Mahan,S.M. and Barbet,A.F.	
AUTHORS		The map1 Gene of Cowdria ruminantium Is a Member of a Multigene	
TITLE		Family Containing Both Conserved and Variable Genes	
JOURNAL		Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)	
REFERENCE		2 (bases 1 to 3507)	
AUTHORS		Sulsona,C.R., Mahan,S.M. and Barbet,A.F.	
TITLE		Direct Submission	
JOURNAL		Submitted (02-FEB-1999) College of Veterinary Medicine Department	
of Pathobiology, University of Florida, P.O. Box 110880,		Gainesville, FL 32610, USA	
FEATURES		Location/Qualifiers	
source		1..3507	
		/organism="Cowdria ruminantium"	
		/isolate="Highway"	
		/db_xref="taxon:779"	
CDS		948..1796	
		/note="ORF2"	
		/codon_start=1	
		/transl_table=11	
		/product="major antigenic protein 1 like protein"	
		/protein_id="AB026347.1"	
		/db_xref="GI:4589101"	
		/translation="MNYKKILVRSALISLMSFLPYGSFAEPVSSNNIGNAKEGFYI	


```
QY 274 TyrGlyGlySerIleGlyIleLysPhe 282
Db 1804 TTTGGTGGTGAACCTGGAGTAAGGTTT 1830
RESULT 12
AX042314 AX042314 843 bp DNA linear PAT 23-NOV-2000
LOCUS Sequence 10 from Patent WO0065063.
DEFINITION AX042314
ACCESSION AX042314
VERSION AX042314.1 GI:11340974
KEYWORDS
SOURCE Ehrlichia chaffeensis.
ORGANISM Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichiae; Ehrlichia; canis group.
REFERENCE 1 (bases 1 to 843)
AUTHORS Barbet,A.F., Bowie,M.V., Cantu,R.R., Burrigge,M.J., Mahan,S.M.,
Whitmire,W.W. and Alleman,A.R.
TITLE Nucleic acid vaccines against rickettsial diseases and methods of
use
JOURNAL Patent: WO 0065063-A 10 02-NOV-2000;
UNIVERSITY OF FLORIDA (US)
FEATURES
source Location/Qualifiers
1..843
/db_xref="taxon:945"
BASE COUNT 282 a 127 c 159 g 275 t
ORIGIN
Alignment Scores: 1.81e-28 Length: 843
Pred. No.: 411.50 Matches: 103
Score: 52.08% Conservative: 47
Percent Similarity: 35.76% Mismatches: 113
Best Local Similarity: 27.25% Indels: 25
Query Match: 6 Gaps: 11
DB:
US-09-846-808-1 (1-284) x AX042314 (1-843)
QY 7 ArgLysPheValLeu-----TrpValMetLeuIleLeuPheThrProHisIleSer 23
Db 12 AAAAAATTTTATACAACTACATAGTAGTCGTAATGCTCTTCTTACCTGGAAATACA 71
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 72 TTTCTGTAGTCAGTACAGAACCAATCTGTGTGTAATTTCTATATCATCAGTGGGAATAT 131
QY 41 LysProAlaArgGlnHisLeuSerLysLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 132 GTACCAAGTGTTCACATTTTGGCGTATTCCTGCTAAACAGGAAGAAATACACAAATC 191
QY 61 GluValPheGlyLeuLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 192 GGAGTATTGGATTAAACAGATGGGATGGCAGCACAAATATCAAAAATCTCCAGAA 251
QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 252 AATACATTTAAGTTCACAAATTTATTCATTTAAATATGAATAATATCCATTTCTAGTTT 311
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 312 GCAGGAGCTGTGGTTATTAATGAAT---GGTCCAGAAATAGAGTTAGAAATGTCCAT 368
QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 369 GAACATTTGATGTAACAAACACAGGTAATAACTATAAGAAC---GATGCTCACAAATAT 425
QY 137 PheAlaLeuAlaLysGluIleSerGlySerAsnProAlaAsnAspLysTyrVal 156
Db 426 TATGCTTTAACCCATMAC--AGTGGGGGAAGCTAAGCAATGCAGGTGATAGTTGTT 482
QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
```

```
Db 483 TTTCTAAAAAATGAAGGACTACTTGATATATCACTTATGTTGAATGCATGCTATGATGA 542
QY 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 543 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTCAGGTTGCTGACTGAT 596
QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 597 TTAATATCCATGTTTGAAGCTATAAACCCATAAAATTTCTTATCAAGGAAAGTTAGGTTG 656
QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 657 AGTTACTCATAGCCAGCAAGCTCTGTTTTTGTGGTGGACATTTTCATAGGTGATA 716
QY 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 717 GGAATGAATTCAGA-----GATATCTCTCTATGATACCC 752
QY 255 ProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
Db 753 AGTACCTCAACTCTCACAGGTAATCACTTACTATAGTAACACTAAGTGTATGCCACTTT 812
QY 275 GlyGlySerIleGlyIleLysPhe 282
Db 813 GGAGTGGAACTTGGAGGAAGGTTT 836
RESULT 13
AF062761 4883 bp DNA linear BCT 18-JUL-1998
LOCUS Ehrlichia chaffeensis 28 kDa major surface antigen multi-gene
DEFINITION locus, partial sequence.
ACCESSION AF062761
VERSION AF062761.1 GI:3327958
KEYWORDS Ehrlichia chaffeensis.
SOURCE Ehrlichia chaffeensis.
ORGANISM Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichiae; Ehrlichia; canis group.
REFERENCE 1 (bases 1 to 4683)
AUTHORS Reddy,G.R., Sulsona,C.R., Barbet,A.F., Mahan,S.M., Burrigge,M.J.
and Alleman,A.R.
TITLE Molecular characterization of a 28 kDa surface antigen gene family
of the tribe Ehrlichiae
JOURNAL Biochem. Biophys. Res. Commun. 247 (3), 636-643 (1998)
MEDLINE 98321180
REFERENCE 2 (bases 1 to 4683)
AUTHORS Reddy,G.R., Sulsona,C.R., Barbet,A.F., Mahan,S.M., Burrigge,M.J.
and Alleman,A.R.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-1998) Diagnostic Medicine/Pathobiology, College
of Veterinary Medicine, Kansas State University, 1800 Denison Ave,
Manhattan, KS 66506, USA
FEATURES
source Location/Qualifiers
1..4683
/organism="Ehrlichia chaffeensis"
/strain="Arkansas"
/db_xref="taxon:945"
<1--134
/note="similar to Cowdria ruminantium MAP1 gene; ORF 1;
silent copy"
/codon_start=3
/transl_table=11
/evidence=not_experimental
/product="28 kDa major surface antigen-1"
/protein_id="AAC26717.1"
/db_xref="GI:3327960"
/translatio="NEFRDISTLKAFATPPSSAATPDLATVTLVSCHFGVGLGRFNF"
179..188
/note="palindromic sequence part 1"
/evidence=not_experimental
203..212
/note="palindromic sequence part 2"
```



```

misc_feature
-35_signal
-10_signal
RBS
CDS

/evidence=not_experimental
226..239
/note="g-rich sequence"
349..354
/evidence=not_experimental
375..380
/evidence=not_experimental
430..434
/evidence=not_experimental
439..1299
/note="similar to Cowdria ruminantium MAP1 gene; ORF 2;
silent copy"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="28 kDa major surface antigen-2"
/db_xref="GI:3327961"
/translation="MNCCKFFITTTALVSLMFLPGISLSDPVQDDNISGNFVYSKYM
PSASHFGVFSKAEERNVTGVFGIQBDWDRCVISRTLSIDFTVPNYSFKYENNPFLG
FAGAIQYSDGPRIELEVSYEAFDVKNQGNKYKNAHRYALSHLLGTETQIDGAGSA
SVFLNELGDKSFMNACDYVISEGIPFSPYICAGIGIDLVSMEFANPKISYOGKLSY
GLSPISPEASVFIGGHFKVIGNEFRDIPITPMSALAGNIPAIIVLDFVFIGI
ELGGRNFQL"
1300..1309
/note="palindromic sequence part 1"
/evidence=not_experimental
1324..1333
/note="palindromic sequence part 2"
/evidence=not_experimental
1347..1361
/note="g-rich sequence"
1473..1479
/evidence=not_experimental
1498..1504
/evidence=not_experimental
1554..1558
/evidence=not_experimental
1563..2399
/note="similar to Cowdria ruminantium MAP1 gene; ORF 3;
silent copy"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="28 kDa major surface antigen-3"
/db_xref="GI:3327962"
/translation="MNCCKFFITTTALVSLMFLPGISLSDPVQDDNISGNFVYSKYM
PSASHFGVFSKAEERNPTVALYGLQDWEGISSSSHNDNHNKNSFKYENNPFLGF
AGAIGYSMDGPRIELEVSYETFDVKNQGNKYKNAHRYALSHLLGTETQIDGAGSA
KSEGLDLSFMNACDYIINESIPLSPYICAGVGTDLISMFEATNPKISYOGKLGSLY
SINPASVFIGGHFKVIGNEFRDIPITLKAFVTSSTPDLAIVLTSVCHFGIELGRRF
NP"
2444..2456
/note="palindromic sequence part 1"
/evidence=not_experimental
2465..2477
/note="palindromic sequence part 2"
/evidence=not_experimental
2491..2505
/note="g-rich sequence"
2618..2624
/evidence=not_experimental
2644..2649
/evidence=not_experimental
2699..2703
/evidence=not_experimental
2708..3350
/note="similar to Cowdria ruminantium MAP1 gene; ORF 4;
silent copy"
/codon_start=1
/transl_table=11
/evidence=not_experimental

```

```

/product="28 kDa major surface antigen-4"
/protein_id="AAC26720.1"
/db_xref="GI:3327963"
/translation="MNCCKFFITTTLVSLMFLPGISLSDPVQDDNISGNFVYSKYM
PSVSHFGVFSKAEERNVTIGVFLQDWDGSTISKNSPENTENVPNYSFKYENNPFLG
FAGAVGYLMNGPRIELEMYSYETFDVKNQGNKYKNAHRYALSHLLGTETQIDGAGSA
FLKNEGLDLSFMNACDYVISEGIPFSPYICAGVGTDLISMFEATNPKISYOGKLGSL
SYSISPEASVFIGGHFKVIGNEFRDIPITPMSALAGNIPAIIVLTSVCHFGIELGRRF"
3586..3596
/note="palindromic sequence part 1"
/evidence=not_experimental
3610..3622
/note="palindromic sequence part 2"
/evidence=not_experimental
3643..3655
/note="g-rich sequence"
3767..3772
/evidence=not_experimental
3788..3793
/evidence=not_experimental
3847..3851
/evidence=not_experimental
3856..>4683
/note="similar to Cowdria ruminantium MAP1 gene; ORF 5;
expressed copy"
/codon_start=1
/transl_table=11
/product="28 kDa major surface antigen-5"
/protein_id="AAC26716.1"
/db_xref="GI:3327959"
/translation="MNYKKVFTTSLISLPLGVSFSDPAGSGINGNFYISGKYMP
SASHFGVFSKAEERNVTGVFLKQNDGSAISNSPNDVFTVSNYSFKYENNPFLGF
AGAIGYSMDGPRIELEVSYETFDVKNQGNKYKNAHRYALSHLLGTETQIDGAGSA
LKNEGLDLSFMNACDYVVGEGIPFSPYICAGIGIDLVSMEATNPKISYOGKLGSL
YSISPEASVFIGGHFKVIGNEFRDIPITPITGSTLAGKNYPAIIVLDFVCHFGIEGM"
BASE COUNT 1578 a 739 c 833 g 1533 t
ORIGIN

Alignment Scores:
Pred. No.: 1.46e-27 Length: 4683
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 1 Gaps: 11

US-09-846-808-1 (1-284) x AF062761 (1-4683)
Qy 7 AtgLySpHeValLeu-----TtpValMetLeuIleLeuPheThrProHisIleSer 23
Db 2717 AAAAAATTTTTTATACAACTACATTAGTATCGCTAATGCTTCTTACTCGGAATATCA 2776
Qy 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyValGlyIleGlnTyr 40
Db 2777 TTTTCTGATCGACGTACAGAACGACAATGTGGTGGTAATTTCTATATACAGTGGGAATAT 2836
Qy 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 2837 GTACCAAGTGTTTTCACATTTTGGCGTATTCCTGCTAAACAGGAAAGAAATACACAATC 2896
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 2897 GGAGTATTTGGATTAAAGCAAGATTTGGGATGGCAGCACATATCTAAAAAATTTCTCCAGAA 2956
Qy 81 ThrAsnPheAsnIle---LysTyAsnProTyTyTyGluAsnAsnArg---LeuGlyPhe 98
Db 2957 AATACATTTAACGTTCCAAATTTATTCATTTAAATATGAAATAATCCATCTTAGGTTTT 3016
Qy 99 SerGlyIlePheGlyTyTyTyTyTyAsnLysAsnPheArgIleGluSerGluLeuSerTyTy 118
Db 3017 GCAGGAGCTGTTGGTTATTATTAATGAAT---GGTCCAGAAATAGAGTTAGAAATGTCCTAT 3073

```


REFERENCE 2 (bases 1 to 1278)
 AUTHORS Reddy, G.R.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAR-1996) Roman G. Reddy, Pathobiology, University of Florida, 471 Mowry Rd., Gainesville, FL 32611-0880, USA

FEATURES
 source Location/Qualifiers

1..1278
 /organism="Cowdria ruminantium"
 /strain="Gardel"
 /db_xref="taxon:779"
 396..1268
 /gene="map1"
 396..1268
 /gene="map1"
 /note="surface protein"
 /codon_start=1
 /transl_table=11
 /product="major antigenic protein 1"
 /protein_id="AAC44143.1"
 /db_xref="GI:1418254"
 /translation="MCKKIFITSLISLVLPVGSFSDVIOEDSNPAGSVVISAKY
 MPTASHFGKMSIKEDSKNTQTVFGLKDDGVKTPSSDSCNNSIIFTEKDYSFKYNN
 PFLGPAAGYSGMNGPRIEFESYFEDVKNGKYNKNDAMHCALDTGTPGSTQAT
 LNSVVKNNLFDIALMLNACVDITLEGMPVSPYVCGAGIGTDLVSVINATNPKLSTQ
 KGLGISYINPEASIFIGGFHFRVIGNEFKDITTSKIFTSTGKLATAASPGFASATLD
 VCHFGIEIGGRFVF"
 BASE COUNT 433 a 174 c 206 g 465 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,74e-26 Length: 1278
 Score: 392.00 Matches: 97
 Percent Similarity: 52.28% Conservative: 52
 Best Local Similarity: 34.04% Mismatches: 114
 Query Match: 25.96% Indels: 22
 DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x CRU50832 (1-1278)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
 Db 432 ATATCATAGTGTGCTTTTACCGGGTGTGCTTTTGTGATGAATACAGAGATAGC 491
 QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
 Db 492 AGCCACGAGTAGTGTATACATAGCGCAAAATACATGCCAATGCATCATCTTTGGT 551
 QY 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
 Db 552 AAAATGTCAATCAAGAAGATTTCAAAAATACTCAAAACAGTATTTGGCCTAAAAAAGAT 611
 QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle----- 85
 Db 612 ---TGGGATGCGTTAAACACCATCATCATAGTAGCGGTAAACAATAGTATTCTTCACT 668
 QY 86 -----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePhe 102
 Db 669 GAAAAGACATATTCAATCAATATGAACACACCATTTTGTAGTTTGTGGAGCAATT 728
 QY 103 GlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHis 122
 Db 729 GGGTATTCAATGAAT---GGTCCAAGAATAAGATTGAAGTATCTTATGAACATTTTGTAT 785
 QY 123 IleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
 Db 786 GTAAAAACCCAGGTGTAATATATAAAAT---GATGCACATATGATGTCCTAGAT 842
 QY 141 LysGluIleSerGlyLysSerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
 Db 843 ACAGGAACACGAGTACTACTACGGGTGCAACATTAATTCATCTGTTATGGTAAAAAT 902
 QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal-----AspGly 178
 Db 178

Db 903 GAAATTTAACTGATATTGCACATAATGTTAAATGCATGCTATGATATAACTTCAAGGA 962
 QY 179 LeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPhe 198
 Db 963 ATGCCA-----GTTTCTCCATATGTGTGTCAGGCATTTGTAGTTAGTGTCAATA 1016
 QY 199 LeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrVal 218
 Db 1017 ATCAATGCTCAAAATCCTAAGTTATCTTATCAGGGAAGTTTAGGTATTAGTTATTCAATA 1076
 QY 219 SerProGlnValSerValPheIleGluGlyTyrHisGlyLeuPheGlyLysLysPhe 238
 Db 1077 AATCCTGAAGCTTCTATCTTTATTGGTGGACATTTCCATAGATATAGTAATGAATTT 1136
 QY 239 GluLysIleProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLys 258
 Db 1137 AAACACATTAATCTACTTCCAAAATATTACCTCAACTGCTAAATTTAGCTACTGCAGCTAGC 1196
 QY 259 ProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIle 278
 Db 1197 CCAGGTTTT-----GCATCAGCAACACTTGATGTTGCCATTTCCGTATAGAAATT 1247
 QY 279 GlyIleLysPheIle 283
 Db 1248 GGAGGAAGGTTTGTA 1262
 RESULT 17
 CRU49843
 LOCUS Cowdria ruminantium major antigenic surface protein (map1) gene, 1564 bp DNA linear BCT 01-MAY-1996
 DEFINITION complete cds.
 ACCESSION U49843
 VERSION U49843.1 GI:1292953
 KEYWORDS heartwater rickettsia strain-Welgevonden stock.
 SOURCE Cowdria ruminantium
 ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 Rickettsiaceae; Ehrlichieae; Cowdria.
 REFERENCE 1 (bases 1 to 1564)
 AUTHORS Fehrsen, J., de Villiers, E.P., Brayton, K.A., van Kleef, M. and Allsopp, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1996) Etienne P. de Villiers, Molecular Biology, Onderstepoort Veterinary Institute, Onderstepoort 0110, South Africa
 FEATURES
 source Location/Qualifiers
 1..1564
 /organism="Cowdria ruminantium"
 /strain="Welgevonden stock"
 /db_xref="taxon:779"
 610..1482
 /gene="map1"
 610..1482
 /gene="map1"
 /note="MAP1; Cr32"
 /codon_start=1
 /transl_table=11
 /product="major antigenic surface protein"
 /protein_id="AAA98610.1"
 /db_xref="GI:1292954"
 /translation="MCKKIFITSLISLVLPVGSFSDVIOEDSNPAGSVVISAKY
 MPTASHFGKMSIKEDSKNTQTVFGLKDDGVKTPSSDSCNNSIIFTEKDYSFKYNN
 PFLGPAAGYSGMNGPRIEFESYFEDVKNGKYNKNDAMHCALDTAOCASATNGAT
 LASSVVKNNLFDIALMLNACVDITLEGMPVSPYVCGAGIGTDLVSVINATNPKLSTQ
 KGLGISYINPEASIFIGGFHFRVIGNEFKDITTSKIFTATNKVSTVANPGFASATLD
 VCHFGIEIGGRFVF"
 BASE COUNT 525 a 216 c 251 g 572 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,75e-26 Length: 1564
 Score: 391.00 Matches: 100
 Percent Similarity: 50.87% Conservative: 47

Best Local Similarity:	34.60%	Mismatches:	112
Query Match:	25.89%	Indels:	30
DB:	1	Gaps:	11
US-09-846-808-1 (1-284) x CRU49843 (1-1564)			
QY	13	ValMetLeuLeuLeuPheThrProHisIleSerLeuAlaSerValLeu	28
DB	646	ATATCATTTAGTGTCAATTTTACCTGGTGTCTCTTCTGATCAATACAGGAGGATAGC	705
QY	29	AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer	48
DB	706	AACCCAGCTGGTGTCTTATCATTCGCAAAATATACCTGCTTACACACTTTGGT	765
QY	49	LysLeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp	68
DB	766	AAATGTCAATTAAGAAGATCAAAAATATCTCAACAGTGTGTGCTTAAAAAAGAT	825
QY	69	LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPheAsn	84
DB	826	-----TGGGATGGAGTTAAAGTACCAACATCAGAAACACCACTTACTCT	870
QY	85	Ile-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe	98
DB	871	TCATCTTTTACTGAAAAGATTTATCTTTCAGATATGAAAACAATCCGTTTTAGGTTTT	930
QY	99	SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118
DB	931	GCTGGAGCAATGGGTACTCAATGAAT---GGACCAAGAAATGAGTTCGAAGTATCCTAT	987
QY	119	GluThrPheHisLysLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis	136
DB	988	GAACACTTTTGATGTAAAAAACCCAGCTGGCACTACAAAAAT---GATCCACACATGAT	1044
QY	137	PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal	156
DB	1045	TGTGCTTTAGATACACGACCAACAGAGTGTCTAATGTTGTCACATTTAGCTTCTGTT	1104
QY	157	ThrLeuLeuAsnAsnGlyIleSerLeuThrSerAlaLeuLeuAsnValCysTyrAspVal	176
DB	1105	ATGATAAAAAATGAAATTTTACAATATATCATTAATGTTTAAATCGGTATGATATA	1164
QY	177	-----AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAsp	194
DB	1165	ATGCTTGTGATGGAATGCCA-----GTTTCTCCATATGTATGTCAGGTATTTGGTACTGAT	1218
QY	195	ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla	214
DB	1219	TTAGTGTCAAGTAATTAATCTCAAAATCTCAATATATCTTATCAAGGAAAGCTAGGCATA	1278
QY	215	SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe	234
DB	1279	AGTTACTCAATCAATCTCGAGCTTCTATCTTATCGGTGGACATTTCCATAGAGTTATA	1338
QY	235	GlyLysLysPheGlyLysIleProValAsnTyrProCysAspTyrProSerProThrPro	254
DB	1339	GGTAATGAATTTAAGATATTTGCTACTTTTAAANAATTTTCCGCACTAATAAAGTATCT	1398
QY	255	ProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyr	274
DB	1399	ACTGTAGCCCAATCCAGGTTTT-----GCATCAGCAACACTTGTATGTTGTCACTTT	1449
QY	275	GlyGlySerIleGlyIleLysPheIle	283
DB	1450	GGTATAGAAATTTGGAGGAGGTTTATA	1476
RESULT 18			
AF368001			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
AF368001.1 GI:15429317			
873 bp DNA linear BCT 06-NOV-2001			
COWDRIA RUMINANTIIUM ISOLATE BURKINA FASO MAJOR ANTIGENIC PROTEIN			
MAPPING			
AF368001			
AF368001.1 GI:15429317			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
BASE COUNT			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-846-808-1 (1-284) x AF368001 (1-873)			
QY	13	ValMetLeuLeuLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn	32
DB	37	ATATCATTTAGTGTCAATTTTACCTGGTGTCTCTTCTGATCAATACAGGAGGATAGC	96
QY	33	-----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer	48
DB	97	ACCCAGCGGCGGAGTGTTCATTTACATTCGCCAAATATACATCCCACTGTCATCATTTTGGT	156
QY	49	LysLeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp	68
DB	157	AAATGTCAATCAAGAAGATCAAAAATATCTCAACAGTATTTGGTCTTAAAAAAGAT	216
QY	69	LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsn-----	82
DB	217	-----TGGGATGGCGTTAAAGTACCAACATCAGAAACACAGTAACAACTTCT	261
QY	83	-----PheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe	98
DB	262	ACAATTTTACTGAAAAGATTTATCTTTCAGATATGAAAACAATCCGTTTTAGGTTTT	321
QY	99	SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118

```
Db 322 GCTGACGAATTGGATACTCAATGAAT---GGGCCAAGAATAGAGTTTGAAGTATCCTAT 378
Qy 119 GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis 136
Db 379 GAAACTTTTGATGTCAAAATCCAGGTGGCACTATAAATAAT-----GATGCCAC 429
Qy 137 PheAlaLeuAlaGlyGluIleSerGlyCysSerGlyCysSerGlyCysSerGlyCysSerVal 156
Db 430 ATGTAATGCTGCTAGACACAGCA-----CAGCCCACTAGTAAATCAAGGGCA 477
Qy 157 ThrLeuIle-----AsnAsnGlyIleSerLeuThrSerAlaLeuIle 170
Db 478 ACATTAGCTTCATCTGTTATGTTAAACAGCAAAATTTAACAGATATATATTAATGTTA 537
Qy 171 AsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIleThrTyrSerCys 188
Db 538 AATCATGTTATGATATAATGCTTGTGGAATGCCA-----GTTCTCCATATGATGT 591
Qy 189 LeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyr 208
Db 592 GCAGTATGCTAGTCTGATTTAGTATCATGTAATTAATGCTACAAATCCTAAATTTATCTAT 651
Qy 209 GlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGly 228
Db 652 CAAGGAAAGCTAGCATAAGTTACTCAATCAATCCTGAAGCTTCTATCTTTATTGGTGG 711
Qy 229 TyrThrHisGlyLeuPheGlyLysPheGluLysIleProValAsnTyrProCysAsp 248
Db 712 CATTTCCATAGATATAGGTAATGAATTAAGATATGCTACTCAAAA-----ATA 765
Qy 249 TyrProSerProThrProAsnSerLysProHisValHisThrThrAlaLeuAlaMet 268
Db 766 TTTACTTCAACTAGTAAGCATCATCTATACCTAAT---CCTGGCTTTGCATCAGCAACA 822
Qy 269 LeuSerIleGlyTyrTyrGlySerIleGlyLysPheIle 283
Db 823 CTTGATGCTGCCATTCGGTATAGAAATTGGAGGAAGGTTTGTGA 867

RESULT 19
AF368007 828 bp DNA linear BCT 06-NOV-2001
LOCUS Cowdria ruminantium isolate Mali major antigenic protein MAP1 gene,
partial cds.
ACCESSION AF368007
VERSION AF368007.1 GI:15429325
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE 1 (bases 1 to 828)
AUTHORS Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaid,A.,
Haydon,D.T., van Heerden,H. and Allsopp,B.A.
TITLE Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants
Are Not Geographically Constrained and Show No Evidence of Having
Evolved under Positive Selection Pressure
J. Clin. Microbiol. 39 (11), 4200-4203 (2001)
JOURNAL 21539003
MEDLINE 11682561
PUBMED
REFERENCE 2 (bases 1 to 828)
AUTHORS Allsopp,M.T., Dorfling,C., Maillard,J.-C., Bensaid,A., Van
Heerden,H. and Allsopp,B.A.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary
Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,
South Africa
FEATURES
source Location/Qualifiers
1 .828
/organism="Cowdria ruminantium"
/isolate="Mali"
/db_xref="taxon:779"
<1..>828
/codon_start=1
CDS
```

```
/transl_table=11
/product="major antigenic protein MAP1"
/protein_id="AAK98147.1"
/db_xref="GI:15429326"
/translation="FISTLISLVFLPGVSDVIQEDNSPVGSVVISAKYMP7ASH
FGMSIKEDSKNTOTVFLKKDWDGKVTSPDSNHSITFEKDYISKRYENPFLGFAG
AIGYSNMGPRIEFESYETFDVKNPGNYKNDAMMYCALDTAAQNSTNGAGLTSMVY
KNLNTDLSLMLNACYDILMDGMPVSPVCAGICTDLVSVINATNPKLSYQGLGISY
SINPEASIFIGGHFHRVIGNEFKDIATSKIFTSKTGISIPNPGFASALDVCHEGIEV
"
BASE COUNT 292 a 141 c 142 g 253 t
ORIGIN
Alignment Scores:
Pred. No.: 1,92e-26 Length: 828
Score: 389.00 Matches: 91
Percent Similarity: 54.07% Conservative: 42
Best Local Similarity: 36.99% Mismatches: 87
Query Match: 25.76% Indels: 26
DB: 1 Gaps: 10
US-09-846-808-1 (1-284) x AF368007 (1-828)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 19 ATATCATAGTGTCTATCTTACCTGGTGTCTTTTCTGATGTAAACAAGAATAAC 78
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 79 AGCCCACTAGTAGTACGTTTACATTAGTGCAAAGTACATGCCCACTGCTTCACACTTTGGT 138
Qy 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 139 AAAATGTCATTAAGGAAGATTCAAGAAATFACCTCAACACAGTATTTGGCTTAATAAAGAT 198
Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPheAsnIle 85
Db 199 -----TGGGATGGGTTTAAACACCATCAGATCTTAGCAATCATCTTCTACA 243
Qy 86 -----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSer 99
Db 244 ATTTTCTAGTAAAGATATTTCTTCAATATGAAACAATCCCTTTTATAGGTTTCGCT 303
Qy 100 GlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGlu 119
Db 304 GGACCAATTGGATACTCAATGAAT---GGACCAAGAATAGAGTTTGAAGTATCTCTATGAA 360
Qy 120 ThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPhe 137
Db 361 ACTTTTGTATGTCAAAAATCCAGCGGGTAATTTATAAACAAC---GATGCACACATGCTACTGT 417
Qy 138 AlaLeuAlaLysGluIleSerGlyCysSerAsnAsnProAlaAsnAsnLysTyrValThr 157
Db 418 GCTTTAGATACAGCAGCAACAAACAGCACTAATGGTCGAGGATTAACATACATCTGTTATG 477
Qy 158 LeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal--- 176
Db 478 GTAAAAACGAAATTTAACAGATATATCACTAATGTTAAATGCGTCTATGATATAATG 537
Qy 177 ---AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThr 195
Db 538 CTTGATGGAATGCCA-----GTCTCCCATATGATGTATGTCAGGATTTTGGCACTGATCTA 591
Qy 196 IleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSer 215
Db 592 GTGTCAGTAATTAATGTCACAAATCCCTAACTATCTTACCAGGCACTTAGTATTAGG 651
Qy 216 TyrThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPheGly 235
Db 652 TATTCAATTAATCCAGAGCATCTATATTTATTGGAGGACACTTTTCATAGAGTTATAGGT 711
Qy 236 LysLysPheGluLysIle 241
:::|||||:::
|||
```



```

/codon_start=1
/product="major antigenic protein 1"
/protein_id="AAC44146.1"
/db_xref="GI:1418260"
/translation="MNCKKIFITSTLISLVSLPGVSFSDVIOEDSNPAGSVVISAKY
MPTASHFGKMSIKEDSRDTPKAVGLKDDVKTSPSSDGSNSIIFTEKDYSPKYENN
PLGFAAGYSGMNGPRIEFVSYETFDVKNPGNYKNDAMKALDTCGTGPGSTOGAT
LNSVMVKNENLTDIALMLNACYDITLEMPYSPYVCAGIGTDLVSVINATPKLSYQ
KGLGISYNPEASIFIGGHRVHVGNEFKDITTSKIFTSTGKLATAASPGFASATLD
VCHGIEIGGRVF"

```

BASE COUNT 431 a 175 c 207 g 465 t
ORIGIN

Alignment Scores:
Pred. No.: 4,01e-26 Length: 1278
Score: 388.00 Matches: 97
Percent Similarity: 51.93% Conservative: 51
Best Local Similarity: 34.04% Mismatches: 115
Query Match: 25.70% Indels: 22
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x CRU50835 (1-1278)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
Db 432 ATATCATTAGTGCATTTTACCTGGTGTCTCTCTTCTCATGTATATACAGGAGTAGC 491
QY 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 492 AACCCAGCGGTAGTGTTCATATAGCGCAAAATACATGCCACTGCTTCACATTTTGGT 551
QY 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 552 AATGTCTCATCAAGAGAGATTCAAAAATACCTCAACACAGTATTTGGCCTAAAAAAGAT 611
QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle----- 85
Db 612 ---TGGATGGCGTTAAACACACCATCATCAGATGCGGTAGCAATAGTATATCTTCACT 668
QY 86 -----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePhe 102
Db 669 GAAAAGACTATTTCATCAATATGAACACACCCATTTTATGGTTTGTCTGGAGCAATT 728
QY 103 GlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHis 122
Db 729 GGGTATTCAATGAAT---GGTCCAGAAGATAGAAATTTGAAGTCTATGAACATTTTGT 785
QY 123 IleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
Db 786 GTAAAAAACCCAGGTGGTAAATATAAAAT---GATGCACATATGATTTGTGCTCTAGAT 842
QY 141 LysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
Db 843 ACAGGAACACCGAGTCTACTCAGGTGGCAACATTAATTCATCTGTATGGTAAATAAAT 902
QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal-----AspGly 178
Db 903 GAAATTTAATCTGATTTGCACTAAATGTTAAATGCATGCTATATACACTTGAGGA 962
QY 179 LeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPhe 198
Db 963 ATGCCA-----GTTTCTCCATATGTTGTGTCAGGCATTTGCTAGTTAGTGTACGTA 1016
QY 199 LeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrVal 218
Db 1017 ATCAATGCTACAATCTCAAGTTATCTTATCAGGGAAGTTAGGTATTTAGTTATCAATA 1076
QY 219 SerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPhe 238
Db 1077 AATCCCTCAAGCTTCTATCTTTATTTGGTGGACATTTCCATAGAGTTATAGTATGAATTT 1136
QY 239 GluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLys 258

```

Db 1137 AAGACATTACTACTTCCAAAATATTACCTCAACTGGTAATATTAGCTACTGCAGTAGC 1196
QY 259 ProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIle 278
Db 1197 CCAGGTTT-----GCATCAACCAACTTGTATGTTTGCATTTCCGTATAGAAATT 1247
QY 279 GlyIleLysPheIle 283
Db 1248 GGAGGAAGGTTTGTGA 1262

RESULT 22
CRMAPI
LOCUS CRMAPI 1467 bp DNA linear BCT 10-OCT-1994
DEFINITION C.ruminantium map1 gene.
ACCESSION X74250
VERSION X74250.1 GI:454266
KEYWORDS immunodominant protein; major antigen; map1 gene; outer membrane protein.
SOURCE heartwater rickettsia.
ORGANISM Cowdria ruminantium
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichiae; Cowdria.
REFERENCE 1 (bases 1 to 1467)
AUTHORS Van Vliet,A.H.M.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-1993) A.H.M. Van Vliet, Inst of Infectious Diseases & Immunology, Dept of Bacteriology, School of Veterinary Medicine, Yalelaan 1, PO Box 80.165, 3508 TD Utrecht, NETHERLANDS
REFERENCE 2 (bases 1 to 1467)
AUTHORS Van Vliet,A.H., Jongejan,F., van Kleef,M. and van der Zeijst,B.A.
TITLE Molecular cloning, sequence analysis, and expression of the gene encoding the immunodominant 32-kilodalton protein of Cowdria ruminantium
JOURNAL Infect. Immun. 62 (4), 1451-1456 (1994)
MEDLINE
FEATURES
source
1..1467
/organism="Cowdria ruminantium"
/isolates="Senegal stock"
/db_xref="taxon:779"
481..1335
/gene="map1"
481..1335
/gene="map1"
/codon_start=1
/transl_table=11
/product="major antigenic protein"
/protein_id="CAA52309.1"
/db_xref="GI:454267"
/db_xref="SPTREMBL:Q46327"
/translation="MNCKKIFITSTLISLVSLPGVSFSDVIOENNPVGSVVISAKY
MPTASHFGKMSIKEDSRDTPKAVGLKDDVKTSPSSDGSNSIIFTEKDYSPKYENNPL
GFAAGYSGMNGPRIEFVSYETFDVKNPGNYKNDAMKALDTCGTGPGSTOGATSYM
VKNENLTDIALMLNACYDITLEMPYSPYVCAGIGTDLVSVINATPKLSYQKGLGIS
YSINPEASIFIGGHRVHVGNEFKDITTSKIFTSSGNASSAVSPGFASATLDVCHFGI
EIGGRVF"
BASE COUNT 493 a 204 c 242 g 528 t
ORIGIN

```

Alignment Scores:
Pred. No.: 4,74e-26 Length: 1467
Score: 388.00 Matches: 100
Percent Similarity: 52.10% Conservative: 49
Best Local Similarity: 34.97% Mismatches: 107
Query Match: 25.70% Indels: 30
DB: 1 Gaps: 13

US-09-846-808-1 (1-284) x CRMAPI (1-1467)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 517 ATATCATTAGTGCATTTTACCTGGTGTGTCATTTTCTGATGTAAATACAGGAAGAAC 576


```
QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 577 AATCCAGTAGGTAGTGTATATATAGCCCAAAATACATGCCCACTGCATCATATTTGCT 636
QY 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 637 AAAATGTCATTAAGAGAGATCTCTAGAGATACATAAGCAGTATTTGGCTTAAAGAAAT 696
QY 69 LeuLeuAsnAspLeuThrGlyIleLys-----AspAsnThrAsn-----Phe 83
Db 697 -----TGGGATGGAGTTTAAACACCACTCGGTAACCAATTCATTTT 741
QY 84 AsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIle 101
Db 742 ACTGAAAAAGAGATATCTTCAAAATGAATAATATCCATTTTATAGTTTTCAGGAGCA 801
QY 102 PheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPhe 121
Db 802 GTTGGATCTCAATGAAT---GGACCAAGATAGAAATTTGAAGTATCTTATGAAACTTTC 858
QY 122 HisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeu 139
Db 859 GAGCTAAGAAATCCAGCGGTAACTACAAAAC-----GATGCCATATGTTTCT 909
QY 140 AlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIle 159
Db 910 GCTCTAGATACAGCTCAAGCTCTACTGCAGGAGCAACTACATCT---GTTATGCTAAAA 966
QY 160 AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal-----Asp 177
Db 967 AATGAAATTTAACAGATATTTTCAATTAATGCTTAAACGCGATGTTATGACATAATGCTGAC 1026
QY 178 GlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp 197
Db 1027 GGAATGCCA-----GTTTCTCCATATGTATGTGAGGCAATGGTACTGACTGTATATCA 1080
QY 198 PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr 217
Db 1081 GTAATTAAGCGTACAAATCTTAATATCTTATCAAGGAAATTAGGAATAAGTTATTCA 1140
QY 218 ValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLys 237
Db 1141 ATAAACCCGGAAGCTCTATCTTTTATTTGGTGGGCACTTCCATAGAGTCTAGGTAACGAA 1200
QY 238 PheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSer 257
Db 1201 TTTAAGATATTCGCTACTCTAAAGTTTCTACTAGAGTGGTAAAGCCAGTAGTCTGTT 1260
QY 258 LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlySer 277
Db 1261 AGTCCAGGTTTT-----GCATCAGCAATACTTGTATGTTTGGCACTTCGGCATAGAA 1311
QY 278 IleGlyIleLysPheIle 283
Db 1312 ATTGGAGGAAGGTTTGTA 1329
RESULT 23
AX042313 AX042313 837 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 9 from Patent WO0065063.
ACCESSION AX042313
VERSION AX042313.1 GI:11340973
KEYWORDS
SOURCE
ORGANISM
Ehrlichia chaffeensis.
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 837)
AUTHORS
Barbet A.F., Bowie M.V., Genta, R.R., Burridge M.J., Mahan, S.M.,
Mcquire, T.C., Rurangirwa, F.R., Moreland, A.L., Simbi, B.H.,
Whitmire, W.W. and Allemen, A.R.
TITLE
Nucleic acid vaccines against rickettsial diseases and methods of
```

```
use
JOURNAL Patent: WO 0065063-A 9 02-NOV-2000;
UNIVERSITY OF FLORIDA (US)
FEATURES Location/Qualifiers
Source 1..837
/organism="Ehrlichia chaffeensis"
/db_xref="taxon:945"
BASE COUNT 279 a 130 c 159 g 269 t
ORIGIN
Alignment Scores: 2.66e-26 Length: 837
Pred. No.: 387.50 Matches: 102
Score: 51.05% Conservative: 44
Best Local Similarity: 35.66% Mismatches: 117
Query Match: 25.66% Indels: 23
DB: 6 Gaps: 10
US-09-846-808-1 (1-284) x AX042313 (1-837)
QY 7 ArgLysPheValLeu-----TrpValMetLeuIleLeuPheThrProHisIleSer 23
Db 12 AAAAAATTTTATACAACTGCATTTAGTAGTATCACTAATGTCCTTTCTACTCGGAATATCA 71
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 72 TTTTCTGATCCAGTGCAGGTGACAAATATTAGTGGTAATTTCTATGTTAGTGCAAGTAT 131
QY 41 LysProAlaArgGluHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 132 ATGCCAAGTGCTTCGATTTTGGCATTTTGGCATGTTTCTGCCAAAGAAAAAATCCTACTGTT 191
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 192 GCATTTGATGGCTTAAACAAGATTGGGAAGGATTAGCTCATCAAGTCACAATGATAAT 251
QY 81 ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 252 ---CATTTCAATAACAAGGGTTATTTCATTTAAATATGAAATAAACCCATTTTATAGGTTT 308
QY 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 309 CGAGGAGCTATTGGTTATTCAATGGGT---GGTCCAAAGAGTAGAGTTTGAAGTGTCTAT 365
QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 366 GAAACATTTGAGCTTAAATAATCAGGTAATACTATAATAAAT---GATGCTCACAGATAC 422
QY 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 423 TGTGCTTTAGGTCAACAGACACACAGCGGAATA-----CCTAAAACTAGTAAATACGTA 476
QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 477 CTGTTAAAAAGCGAAGGATTGCTTGACATCATTTATGCTTAATGCTATGATATA 536
QY 177 AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
Db 537 ATAACAGAGAGCATACCTTTGCTCTCTTACATATGTCAGAGTCTTGGTACTGATTAAATA 596
QY 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 597 TCCATGTTTGAAGCTACAAATCTTAAATTTTCTACCAAGGGAAGTTAGGTCTCTAAGTTAC 656
QY 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
Db 657 TCTATAAACCCAGAGAGCTTCTGTTATTATTGTTGGACATTTTCATAGGTGATAGGAAC 716
QY 237 LysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsn 256
Db 717 GAATTTAGGACATCTCTACTCTGAAGAGCATTTGTTACGTATCATCAGCTACTCA----- 770
QY 257 SerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
```

```

Db 771 -----GATCTAGCAATAGTAACTAGTGTATGTCATTTTGAATA 812
|||||
|||
Db 277 SerileGlyIleLysPhe 282
|||||
|||
Db 813 GAACCTGGAGGAGGTTT 830
|||||
|||

RESULT 24
AF077732
LOCUS
DEFINITION
Ehrlichia chaffeensis strain 91HE17 outer membrane protein p28
precursor, gene, complete cds.
ACCESSION
AF077732
VERSION
AF077732.2 GI:14488429
KEYWORDS
Ehrlichia chaffeensis.
SOURCE
Ehrlichia chaffeensis
ORGANISM
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 1283)
Yu, X.-J., McBride, J.W., and Walker, D.H.
Genetic diversity of the 28-kilodalton outer membrane protein gene
in human isolates of Ehrlichia chaffeensis
J. Clin. Microbiol. 37 (4), 1137-1143 (1999)
MEDLINE
99175287
PUBMED
10074538
REFERENCE
2 (bases 1 to 1283)
Yu, X.-J. and Walker, D.H.
Direct Submission
AUTHORS
Yu, X.-J. and Walker, D.H.
TITLE
Submitted (14-JUL-1998) Pathology, The University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
JOURNAL
3 (bases 1 to 1283)
Yu, X.-J. and Walker, D.H.
Direct Submission
AUTHORS
Yu, X.-J. and Walker, D.H.
TITLE
Submitted (19-JUN-2001) Pathology, The University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
JOURNAL
REMARK
Sequence update by submitter
COMMENT
On Jun 19, 2001 this sequence version replaced gi:3414966.
FEATURES
Location/Qualifiers
source
1..1283
/organism="Ehrlichia chaffeensis"
/strain="91HE17"
/db_xref="taxon:945"
236..1078
/codon_start=1
/transl_table=11
/product="outer membrane protein p28 precursor"
/protein_id="AAC31545.1"
/db_xref="GI:3414967"
/translation="MNYKKVFITTSALISLIPGVSFSDPTGSGINGNFYISGKYP
SASHGVFSKAEERSTAGVGLKODMGSAISHTTPENLFTVSNYSFKYENPFLGF
AGAIGYMDGPRIELVETEDVKNQGSYKNEAHRYCALSHHSGTSMSSAGDFV
FLKNEGLLDVSMFNACYDVISGIPFSCVIGTDLVSMFEATNPKNPKISYOGKGLG
SYSISPTSVFICGGHFHKVIGNEFRDIPAVIPSGSLTGNHPAIVLTDVCHFGIELGG
REF"
BASE COUNT 411 a 204 c 242 g 426 t
ORIGIN

Alignment Scores:
Pred. No.: 6.12e-26 Length: 1283
Score: 386.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best local Similarity: 34.17% Mismatches: 115
Query Match: 25.56% Indels: 20
DB: Gaps: 9

US-09-846-808-1 (1-284) x AF077732 (1-1283)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
::: ||||| ||| ::::| :::
Db 272 ATATCACTAATATCTTCTCTACCTGGAGTAGTATCTTCTGACCCACAGCAGTAGTGGTATT 331
::: ||||| ||| ::::| :::
QY 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
::: ||||| ||| ::::| :::

```

Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA

FEATURES

Source
1. .1307
Location/Qualifiers
/organism="Ehrlichia chaffeensis"
/strain="V8"
/db_xref="taxon:945"
236. .1078
/note="p28-19; 28 kDa outer membrane protein"
/codon_start=1
/transl_table=11
/product="outer membrane protein p28"
/protein_id="AAL12924.1"
/db_xref="GI:15991534"
/translation="MNKKVFITSLISLPGVFSFSDPTGSGINGNFYISGKYP
SASHFGVSAKEERSTTAGVGLKODWGSALSHHTPENITVSNYSFKYNNPFLGF
AGAIGYMDGPRIELEVSYETFDVKNQNSYKNEAHRYCALSHHSGTSMSSAGDKFV
FLKNEGLDVSFMLNACYDVISEGIPIFYICVIGTDLVSMFEATNPKISVQGLGL
SYSISPETSVFVIGGFHFKVIGNEFRDIPAVIPSGSSLTGNHPAIVLDDVCHFGIELGG
RFAF"

BASE COUNT 420 a 212 c 245 g 430 t

ORIGIN

Alignment Scores:
Pred. No.: 6,25e-26 Length: 1307
Score: 386.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.56% Indels: 20
DB: 1 Gaps: 9

US-09-846-808-1 (1-284) x AF393394 (1-1307)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
Db 272 ATATCACTAATCTCTCTACCTGGAGTATCATTTCTGACCAACACAGTAGTGAT 331
QY 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 332 AACGGCAATTTCTATATCATGCTGAAAAATATATGCCAAGCGCTTCGCATTTGGGGTGT 391
QY 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 392 TCTGCTAAGGAAGAAAGAGTACACAGCTGGAGTATTTGGATTGAAGCAAGATTGGGAT 451
QY 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 452 GCGAGTGCAATATCTCACACCACCCAGAAATATATTCACGTCTTCAATATTCTGTT 511
QY 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLys 108
Db 512 AAATATGAAATAAACCCATTTTAGGATTTCGAGGAGCTATTGGTTACTCAATGGAT--- 568
QY 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 569 GCGCCCAAGAAATAGAGCTTGAAGTATCTTATGACACATTCGATGTAAATAATCAAGTAAC 628
QY 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 629 AGTTACAAGAAC---GAACGGCAGTAGTATTGTGCTTTATCATCATCAGTTCAAGAAACA 685
QY 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 686 AGCATCTCAGTGCAGGTATATAATTTGTTTCTAAAAAATGAAGGATTACTTTCAGCTA 745
QY 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184
Db 746 TCGTTTCATGCTGAACGCATGCTATGATGTAATAAGTGAAGGAATACCTTTTCT----- 799
QY 185 ThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThr 204
Db 800 CCTTACATATGTAGTAGTATCGGAGCTGATTTAGTATCCATGTTTGAACCTCAAACTCT 859
QY 205 LysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerVal 224

Db 860 AAAATTCTTACCAAGGAAGTTAGGTTTAAAGCTACTATATAAGCCAGAAACTTCTGTG 919
QY 225 PheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsn 244
Db 920 TTTATTGGTGACATTTTTCATAAGGTAGTAGGAACGAATTTAGA----- 964
QY 245 TyrProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThr 264
Db 965 -----GATATTCTGCTGTAATACCTAGTGGATCAAGTCTTACAGGAACACCCT 1015
QY 265 AlaLeuAlaMetLeuSerIleGlyTyrTyrGlySerIleGlyLysPhe 282
Db 1016 GCAATAGTAACACTGGACGTATGCCACTTTGGTATAGAGCTTGAGGAAGGTTT 1069
RESULT 26
AF393390
LOCUS
DEFINITION
AF393390 . 1309 bp DNA linear BCT 10-OCT-2001
Ehrlichia chaffeensis strain V4 outer membrane protein p28 gene,
complete cds.
VERSION
AF393390.1 GI:15991533
KEYWORDS
Ehrlichia chaffeensis.
SOURCE
Ehrlichia chaffeensis
ORGANISM
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 1309)
AUTHORS
Long,S.W., Zhang,X.-F., Qi,H., Standaert,S., Walker,D.H. and
Yu,X.-J.
TITLE
Allele variation and patterns of transcription of the Ehrlichia
chaffeensis 28 kDa outer membrane protein multigene family
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1309)
AUTHORS
Yu,X.-J., Zhang,X.-F. and Walker,D.H.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUN-2001) Pathology, University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA
FEATURES
Location/Qualifiers
1. .1309
Source
/organism="Ehrlichia chaffeensis"
/strain="V4"
/db_xref="taxon:945"
238. .1080
/note="p28-19; 28 kDa outer membrane protein"
/codon_start=1
/transl_table=11
/product="outer membrane protein p28"
/protein_id="AAL12920.1"
/db_xref="GI:15991534"
/translation="MNKKVFITSLISLPGVFSFSDPTGSGINGNFYISGKYP
SASHFGVSAKEERSTTAGVGLKODWGSALSHHTPENITVSNYSFKYNNPFLGF
AGAIGYMDGPRIELEVSYETFDVKNQNSYKNEAHRYCALSHHSGTSMSSAGDKFV
FLKNEGLDVSFMLNACYDVISEGIPIFYICVIGTDLVSMFEATNPKISVQGLGL
SYSISPETSVFVIGGFHFKVIGNEFRDIPAVIPSGSSLTGNHPAIVLDDVCHFGIELGG
RFAF"

BASE COUNT 420 a 212 c 247 g 430 t

ORIGIN

Alignment Scores:
Pred. No.: 6,27e-26 Length: 1309
Score: 386.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.56% Indels: 20
DB: 1 Gaps: 9

US-09-846-808-1 (1-284) x AF393390 (1-1309)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30

Db 274 ATATCACTAATCTCTCTACCTGGAGTATCATTTTCTGACCAACAGTAGTGATTT 333

Qy 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 334 AACGCAATTTCTATATCAGTGGAAATATATATGCAAGCGCTTGCATTTGGGGTGTTT 393
Qy 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 394 TCTGCTAAGGAAGAAAGATGACACAGCTGGAGTATTGGATTGAACAAGATTGGGAT 453
Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 454 GGCAGTGAATATCTCACACACCCCAAGAAATATATATTCATCTGTTCAAAATATTCGTTT 513
Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
Db 514 AAATATGAAATAACCCATTTTGGAGTATTCGAGGAGCTATTGGTTACTCAATGAT--- 570
Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 571 GGCCCAAGATAGAGCTTGAAGTATCTTATGAGACATTCGATGATAAATAAATCAAGGTAAC 630
Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 631 AGTTACAAGAAC---GAAGCCATAGGTATTGCTTTATCTCATCAGTTCAGGAACA 687
Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 688 AGCATGTCGAGTCGAGGTGATAAATTTCTTTCTTAAATAATGAAGGATTACTTCACGTA 747
Qy 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184
Db 748 TCGTTCATGCTGAACGATGCTATGATGATAAAGTGAAGAAATACCTTTTCT----- 801
Qy 185 ThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThr 204
Db 802 CCTACATATGTTAGGTATCGGACTGATTTAGTATCCAGTTTGAAGCTACAAATCCT 861
Qy 205 LysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerVal 224
Db 862 AAAATTTCTTACCAAGGAAGTATAGGTTTAAGCTACTATATAAGCCCAAGAACTCTCTG 921
Qy 225 PheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsn 244
Db 922 TTTATTTGTGTGACATTTTTCATTAAGGTGATAGGAACGAATTTAGA----- 966
Qy 245 TyrProCysAspTyrProSerProThrProAsnSerLysProHisValHisThrThr 264
Db 967 -----GATATTCCTGCTGATATACCTAGTGGATCAAGTCTTACAGGAACACCCCT 1017
Qy 265 AlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPhe 282
Db 1018 GCAATAGTAACACTGGAGGTATGCCACTTTGGTATAGACTTGGAGGAAGGTTT 1071
RESULT 27
AF077733
LOCUS 1243 bp DNA linear BCT 20-JUN-2001
DEFINITION Ehrlichia chaffeensis strain Jax outer membrane protein p28
ACCESSION AF077733
VERSION AF077733.2 GI:14495545
KEYWORDS Ehrlichia chaffeensis.
SOURCE Ehrlichia chaffeensis
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
1 (bases 1 to 1243)
REFERENCE
AUTHORS Yu,X.-J., McBride,J.W. and Walker,D.H.
TITLE Genetic diversity of the 28-kilodalton outer membrane protein gene
in human isolates of Ehrlichia chaffeensis
J. Clin. Microbiol. 37 (4), 1137-1143 (1999)
JOURNAL 99175287
MEDLINE 10074538
PUBMED
REFERENCE 2 (bases 1 to 1243)
AUTHORS Yu,X.-J. and Walker,D.H.

TITLE Direct Submission
JOURNAL Submitted (14-JUL-1998) Pathology, The University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
REFERENCE 3 (bases 1 to 1243)
AUTHORS Yu,X.-J. and Walker,D.H.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2001) Pathology, The University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
REMARK Sequence update by submitter
COMMENT On Jun 20, 2001 this sequence version replaced gi:3414968.
FEATURES
source
1..1243
/organism="Ehrlichia chaffeensis"
/strain="Jax"
/db_xref="taxon:945"
237..1079
/codon_start=1
/transl_table=11
/product="outer membrane protein p28 precursor"
/protein_id="AAC31546.1"
/db_xref="GI:3414969"
/translation="MNVKKVFITTSALISLISLPQVSFSDPTGSGINGNEFVSGKYM
P
SASHFGVPSAKEERSTTAGVGLQKDWGSAISHTTPTENIFTVSNYSFKYNNPPLGF
AGAIGSMDGPRIELEVSYETFDVKNQNSYKNAHRYCALSHSSGSMAGDKFY
FLKNEGLLDSFMLNACYDVISGIPFYICVIGITLDVSMFATNPKISYQKGL
SYSISPETSVFIGGFHKVIGNEFRDIPAVIPNGSSLTGNHPAIVTLDVCHFGIELGG
RFAP"
BASE COUNT 396 a 197 c 233 g 417 t
ORIGIN
Alignment Scores:
Pred. No.: 7.25e-26 Length: 1243
Score: 385.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.50% Indels: 20
DB: 1 Gaps: 9
US-09-846-808-1 (1-284) x AF077733 (1-1243)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
Db 273 ATATCACAATATCTCTCTACCTGGAGTATCATTTCTGACCCCAACAGGTAGTGTATT 332
Qy 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 333 AACGCAATTTCTATATCAGTGGAAATATATATGCAAGCGCTTGCATTTGGGGTGTTT 392
Qy 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 393 TCTGCTAAGGAAGAAAGATGACACAGCTGGAGTATTGGATTGAACAAGATTGGGAT 452
Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 453 GGCAGTGAATATCTCACACACCCCAAGAAATATATATTCATCTGTTCAAAATATTCGTTT 512
Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
Db 513 AAATATGAAATAACCCATTTTGGAGTATTCGAGGAGCTATTGGTTACTCAATGAT--- 569
Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 570 GGCCCAAGATAGAGCTTGAAGTATCTTATGAGACATTCGATGATAAATAAATCAAGGTAAC 629
Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 630 AGTTACAAGAAC---GAAGCCATAGGTATTGCTTTATCTCATCAGTTCAGGAACA 686
Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 687 AGCATGTCGAGTCGAGGTGATAAATTTCTTTCTTAAATAATGAAGGATTACTTCACGTA 746
Qy 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184

Direct Submission
Submitted (14-JUL-1998) Pathology, The University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
3 (bases 1 to 1243)
Yu, X.-J. and Walker, D.H.

Direct Submission
Submitted (20-JUN-2001) Pathology, The University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
Sequence update by submitter
On Jun 20, 2001 this sequence version replaced gi:3414968.

Location/Qualifiers
1..1243

/organism="Ehrlichia chaffeensis"
/strain="Jax"
/db_xref="taxon:945"
237..1079
/codon_start=1
/transl_table=11
/product="outer membrane protein p28 precursor"
/protein_id="AAC31546.1"
/db_xref="GI:3414969"

/translation="MNVKKVFITTSALISLISLPQVSFSDPTGSGINGNEFVSGKYM
P
SASHFGVPSAKEERSTTAGVGLQKDWGSAISHTTPTENIFTVSNYSFKYNNPPLGF
AGAIGSMDGPRIELEVSYETFDVKNQNSYKNAHRYCALSHSSGSMAGDKFY
FLKNEGLLDSFMLNACYDVISGIPFYICVIGITLDVSMFATNPKISYQKGL
SYSISPETSVFIGGFHKVIGNEFRDIPAVIPNGSSLTGNHPAIVTLDVCHFGIELGG
RFAP"

BASE COUNT 396 a 197 c 233 g 417 t

ORIGIN
Alignment Scores:
Pred. No.: 7.25e-26 Length: 1243
Score: 385.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.50% Indels: 20
DB: 1 Gaps: 9

US-09-846-808-1 (1-284) x AF077733 (1-1243)

Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
Db 273 ATATCACAATATCTCTCTACCTGGAGTATCATTTCTGACCCCAACAGGTAGTGTATT 332

Qy 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 333 AACGCAATTTCTATATCAGTGGAAATATATGCAAGCGCTTGCATTTGGGGTGTTT 392

Qy 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 393 TCTGCTAAGGAAGAAAGATGACACAGCTGGAGTATTGGATTGAACAAGATTGGGAT 452

Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 453 GGCAGTGAATATCTCACACACCCCAAGAAATATATATTCATCTGTTCAAAATATTCGTTT 512

Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
Db 513 AAATATGAAATAACCCATTTTGGAGTATTCGAGGAGCTATTGGTTACTCAATGAT--- 569

Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 570 GGCCCAAGATAGAGCTTGAAGTATCTTATGAGACATTCGATGATAAATAAATCAAGGTAAC 629

Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 630 AGTTACAAGAAC---GAAGCCATAGGTATTGCTTTATCTCATCAGTTCAGGAACA 686

Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 687 AGCATGTCGAGTCGAGGTGATAAATTTCTTTCTTAAATAATGAAGGATTACTTCACGTA 746

Qy 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184

SOURCE	Ehrlichia chaffeensis.	CDS	2498..3352
ORGANISM	Ehrlichia chaffeensis		/gene="omp-1n"
	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;		/note="omp-1 family member"
	Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		/codon_start=1
REFERENCE	1 (bases 14844 to 21136; 21479 to 22234)		/transl_table=11
AUTHORS	Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.		/product="major outer membrane protein omp-1n"
TITLE	Immunodominant major outer membrane proteins of Ehrlichia chaffeensis are encoded by a polymorphic multigene family		/protein_id="AAK28659.1"
JOURNAL	Infect. Immun. 66 (1), 132-139 (1998)		/db_xref="GI:13511830"
MEDLINE	98084465		/translation="MSKRSNRKFEVLWMLILFTPHISLASVLNDHNSMYGVIOYKPAR
PUBMED	9423849		OHLKLLIKESAANTVEVFGKLDLLDLGDKDNTNFENIKYPIYENNRGLGSGIF
	2 (bases 1 to 27190)		GYYNKFRIESELSYEFHKKNGYKIDCKEKFALKEISGGSNNPANNKYVTLIN
REFERENCE	Ohashi,N., Rikihisa,Y. and Unver,A.		NGISLTSALINVCYDVLGKNIITYSCLGFGVDTIDFLSKYTTKFSQGLGASVY
AUTHORS	Analysis of Transcriptionally Active Gene Clusters of Major Outer		SPQSVFIEGYHGLGKFEKIPVNPYPCDYPSTPPNPKPHVHTTALAMLSIGYVG
TITLE	Membrane Protein Multigene Family in Ehrlichia canis and E. chaffeensis		SIGIKFIL"
	3 (bases 1 to 27190)		/codon_start=1
JOURNAL	Infect. Immun. 69 (4), 2083-2091 (2001)		/transl_table=11
MEDLINE	21153566		/product="unknown function UN2"
PUBMED	11254561		/protein_id="AAK28660.1"
REFERENCE	3 (bases 1 to 27190)		/db_xref="GI:13511831"
AUTHORS	Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.		/translation="MNNKKSQGVLLFAAYSIFKYINSCTSKFRLSVKCFVYNPLKYF
TITLE	Direct Submission		YKSTFLKLGKFFVYNYKSTSNVLRSLFKLGLIIRPTNEKISYLNPTNPRVLIY
JOURNAL	Submitted (24-SEP-1996) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	gene	3915..4808
	4 (bases 1 to 27190)		/gene="omp-1q"
REFERENCE	Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.		/codon_start=1
AUTHORS	Direct Submission		/transl_table=11
TITLE	Submitted (27-AUG-1997) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	CDS	3915..4808
JOURNAL	5 (bases 1 to 27190)		/gene="omp-1q"
REFERENCE	Ohashi,N., Rikihisa,Y. and Unver,A.		/note="omp-1 family member"
AUTHORS	Direct Submission		/codon_start=1
TITLE	Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	gene	4825..5682
REMARK	Sequence update by submitter		/gene="omp-1p"
COMMENT	On Apr 2, 2001 this sequence version replaced gi:2853584	CDS	4825..5682
FEATURES	Location/Qualifiers		/gene="omp-1p"
	1..27190		/note="omp-1 family member"
	/organism="Ehrlichia chaffeensis"		/codon_start=1
	/strain="Arkansas"		/transl_table=11
	/db_xref="taxon:945"		/product="major outer membrane protein omp-1p"
	315..956		/protein_id="AAK28662.1"
CDS			/db_xref="GI:13511833"
	/codon_start=1		/translation="MOKLYISFIILSGLLLPKYVFCMHNNIDGSYVTIKYQLTTPH
	/transl_table=11		FKRPIKETSDFDQEPICLAKITANTKEDTLKENFSFSLHQTDSYKSYQNDLLGIGL
	/product="hypothetical transcriptional regulator"		SVGLFVSFRFEFEAGAYNFNKRRLARYKSDGYKFAIPRKSEHGEJLNDTFGTVAK
	/protein_id="AAC02808.2"		NGLSIISNTINLCSETKYKAFPTYICIGVGDFEIFDVMRIKFAVQGVGYVPT
	/db_xref="GI:13511828"		SKLITSINGQYHKVIGNKFE LLPYQPVLPKRLVTNKTSKDIDQDVTAASLTNLEHFS
	/translation="MSTHANKINKMEDTEIKAHDLPIAGLIDKNQKKNFTNKGVN		SEIGLSIF"
	DDTSKFTKEGVESTKDLKNSTTKIRHPVDECVGKEIKRQRMGMSQNLANKLG	gene	5695..6513
	ITFQVQKYKNGYRNIVISRLYQASVLYNEVRDMLKQLEDKNIISCDHTVASTSL		/gene="omp-1t"
	KDNEKEPIPEFHDSDKIDSKVLMVRAVTCIKNEKVRNIINLVKALSLDNQ"		5695..6513
	1136..2032		/gene="omp-1t"
gene			/note="omp-1 family member"
	/gene="omp-1m"		/codon_start=1
CDS	1136..2032		/transl_table=11
	/gene="omp-1m"		/product="major outer membrane protein omp-1m"
	/note="omp-1 family member"		/protein_id="AAK28658.1"
	/codon_start=1		/db_xref="GI:13511829"
	/transl_table=11		/translation="MILINMKLFYHLDNIMTKFSAIGIVLSVLTFLFACNVFASPIPID
	/product="major outer membrane protein omp-1m"		FSNESETAGYASGHYNIQFPSPISVKYKTDETEKELTLFTLKESTDTPTFKOKS
	/protein_id="AAK28658.1"		EFNDKCGSPILYRNNTYGFSGAIGYSGGLRVELEGAFTFDDVKOKYKKNYRYFAL
	/db_xref="GI:13511829"		CKDSDIESTDNCNHNVMKNEGFRVLSLTFNACYDINVSNSLSPSACGIGQGITN
	/translation="MILINMKLFYHLDNIMTKFSAIGIVLSVLTFLFACNVFASPIPID		FLGDTLHTLFLKALGLGFLDISPTKTVIFANGYYVTKDKNSFTNLSVQVPIPLEAKPH
	FSNESETAGYASGHYNIQFPSPISVKYKTDETEKELTLFTLKESTDTPTFKOKS	gene	2498..3352
	EFNDKCGSPILYRNNTYGFSGAIGYSGGLRVELEGAFTFDDVKOKYKKNYRYFAL		/gene="omp-1n"
	CKDSDIESTDNCNHNVMKNEGFRVLSLTFNACYDINVSNSLSPSACGIGQGITN		
	FLGDTLHTLFLKALGLGFLDISPTKTVIFANGYYVTKDKNSFTNLSVQVPIPLEAKPH		
	IDPTACFNADNYGGEVGLRFIL"		
	2498..3352		
gene			
	/gene="omp-1n"		


```
/protein_id="CAC17072.1"
/db_xref="GI:11340966"
translation="MNCKKIFITSTLISLVSFLPGVSPDSVIOEDSNPAGSVYISAKY
MPTASHFGKMSIKEDSKNTOTVFLGKKDMDGVKTPSDSSNTNFTTEKDYSPRENN
PFLGFAIGAIGSMNGPRIEFVSYETFDVKNLGNYNKNAHMYCALDTAAQNSTNGAG
LTTSMVKNEMLTNSLMNACYDMLDGPVSPVCAGIGTDLVSVINATNPKLSYQ
KGLGISYSINSEASIFIGGHPHRVIGNEFKDIATLKIFTSKTGISNPGFASATLDVCH
FGIEIGGRFVF"
BASE COUNT      307 a 139 c 149 g 269 t
ORIGIN
Alignment Scores:
Pred. No.:      6,37e-26      Length:      864
Score:          383.50      Matches:      100
Percent Similarity: 50.17%      Conservative: 46
Best Local Similarity: 34.36%      Mismatches: 108
Query Match:     25.40%      Indels:      37
DB:              6          Gaps:      12

US-09-846-808-1 (1-284) x AX042305 (1-864)
Qy 13 ValMetLeuLeuPheThrProHisValHisSerLeuAlaSerValLeu----- 28
Db 37 ATATCATTAGTGCATTTTACCTGGTGGTCTCTTTCTGATGTAATACAGGAAGACAGC 96
Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 97 AACCCAGCAGGAGTGTATACATTAGCCCAAAATACATGCCCACTGCATCTTTGGT 156
Qy 49 LysLeuLeuIleGlyGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 157 AAAATGTCATCAAAAGAGATCAAAAAATACTCAACGGTATTTGGTCTAAAAAAGAT 216
Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPhe 83
Db 217 -----TGGGATCGCGCTTAAACACCACTCAGATCTAGCAATACTAATCTT 261
Qy 84 AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 262 ACAATTTTACTGAAAGACATATTCTTTCAGATATGAAACAAATCCGTTTTAGTTTC 321
Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 322 GCTGGAGCAATTTGGGTACTCAATGAAT---GGACCAAGAATAAGAGTTTCAAGTATCCTAT 378
Qy 119 GluThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHis 136
Db 379 GAACCTTTTGATGTAATAAACCTAGTGGCAACTATAAAAC---AACGCACACATGTAC 435
Qy 137 PheAlaLeuAlaLysGluIleSerGlyLysSerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 436 TGTGCTTTAGATACACGACGACCAAAATAGCACTAATAGCGCGAGGATTAACTACATCTGT 495
Qy 157 ThrLeuLeuAsnAsnGlyIleSerLeuThrSerAlaLeuLeuAsnValCysTyrAspVal 176
Db 496 ATGGTAAAAACGAAATTTAAACAAATATATCATTAATGTTAAATCCGTTTATGATATC 555
Qy 177 -----AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 556 ATGCTTGATGGAATACCA-----GTTTCTCCATATGATGTCAGGATTATGGCAGTAC 609
Qy 195 ThrIleAspPhePheLysLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 610 TTAGTGTCTAGTAATTAATCTACAAATCCTAAATATCTTATCAAGGAAGCTAGGCATA 669
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 670 AGTTACTCAATCAATCTGAAGCTTCTATCTTTATCGGTGGACATTTCCATAGAGTTATA 729
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 730 GGTAATGAATTTAA-----GATATTGCTACCTTAAATAAATA 765
```

```
Qy 255 ProAsnSerLysProHisValHisThrThr-----AlaLeuAlaMetLeuSerIleGly 272
Db 766 TTACTTCAAAAAACAGGAATATCTAATCTCGCTTGGCATCAGCAACACTTGATTTGT 825
Qy 273 TyrTyrGlyGlySerIleGlyIleLysPheIle 283
Db 826 CACTTTGGTATAGAAATTCGAGGAGGTTTGTGA 858

RESULT 32
CRU50833      1265 bp      DNA      linear      BCT 14-JUL-1996
LOCUS      Cowdria ruminantium Highway major antigenic protein 1 (map1) gene,
DEFINITION      complete cds.
ACCESSION      U50833
VERSION      U50833.1 GI:1418255
KEYWORDS      . heartwater rickettsia strain=Highway.
SOURCE      Cowdria ruminantium
ORGANISM      Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
      Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE      1 (bases 1 to 1265)
AUTHORS      Reddy,G.R., Sulsona,C.R., Harrison,R.H., Mahan,S.M., Burrridge,M.J.
      and Barbet,A.F.
TITLE      Sequence heterogeneity of the major antigenic protein 1 genes from
      Cowdria ruminantium isolates from different geographical areas
      Clin. Diagn. Lab. Immunol. 3 (4), 417-422 (1996)
MEDLINE      96400830
REFERENCE      2 (bases 1 to 1265)
AUTHORS      Reddy,G.R.
DIRECT SUBMISSION
TITLE      Submitted (06-MAR-1996) Roman G. Reddy, Pathobiology, University of
      Florida, 471 Mowry Rd., Gainesville, FL 32611-0880, USA
JOURNAL
FEATURES
      source
          1..1265
              /organism="Cowdria ruminantium"
              /strain="Highway"
              /db_xref="taxon:779"
          383..1246
              /gene="map1"
          383..1246
              /gene="map1"
              /note="surface protein"
              /codon_start=1
              /transl_table=11
              /product="major antigenic protein 1"
              /protein_id="AAC44144.1"
              /db_xref="GI:1418256"
      translation="MNCKKIFITSTLISLVSFLPGVSPDSVIOEDSNPAGSVYISAKY
      MPTASHFGKMSIKEDSKNTOTVFLGKKDMDGVKTPSDSSNTNFTTEKDYSPRENN
      PFLGFAIGAIGSMNGPRIEFVSYETFDVKNLGNYNKNAHMYCALDTAAQNSTNGAG
      LTTSMVKNEMLTNSLMNACYDMLDGPVSPVCAGIGTDLVSVINATNPKLSYQ
      KGLGISYSINSEASIFIGGHPHRVIGNEFKDIATLKIFTSKTGISNPGFASATLDVCH
      FGIEIGGRFVF"
BASE COUNT      444 a 175 c 194 g 452 t
ORIGIN
Alignment Scores:
Pred. No.:      1,01e-25      Length:      1265
Score:          383.50      Matches:      100
Percent Similarity: 50.17%      Conservative: 46
Best Local Similarity: 34.36%      Mismatches: 108
Query Match:     25.40%      Indels:      37
DB:              1          Gaps:      12

US-09-846-808-1 (1-284) x CRU50833 (1-1265)
Qy 13 ValMetLeuLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
Db 419 ATATCATTAGTGCATTTTACCTGGTGGTCTCTTTCTGATGTAATACAGGAAGACAGC 478
Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 479 AACCCAGCAGGAGTGTATACATTAGCCCAAAATACATGCCCACTGCATCTTTGGT 538
```



```
Db 24471 GGTCTTATGAGAAATTTGAACCTGAAAGACAATGGTACCCTGAGATGACCAAGCTTAC 24412
Qy 135 LysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLys 154
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24411 AAATTTTGGTTGCTCGAAATGCTACAAT-----AGTGATATAAG 24367
Qy 155 TyrValThrLeuIleAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyr 174
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24366 TTTATAGTACTAGAGAAATACCGGCTTCTGACAAAGTCTCTTAATGTAAATGTTGTTAT 24307
Qy 175 AspValaspGlyLeuLysHisAsnIleIleThrTyrSerCysLysLeuGlyPheGlyValasp 194
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24306 GATATTGCTAGTGGTAGTATTCCCTTAGCACCTTATATGCTGCTGGTGGTGGTCAGAT 24247
Qy 195 ThrIleaspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24246 TATATAAGTTTGTAGGTATATCATGCTAAGTTTCTTATCAAGTTAAGTTTGGTGTCT 24187
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24186 AACTACCCTCTAAATGTTAATACTATGTTGTTGGTGGGGCTTATTACCATAAGTTGTA 24127
Qy 235 GlyLysLysPheGluLysIleProValasnTyrProCysAspTyrProSerProThrPro 254
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24126 GGTATAGGTATGAGAGAGTAGAATAAGCTTAC-----CATCCTACTGCA 24082
Qy 255 ProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24081 TTATCTGAGTCTCTAGAACT---ACTTCAGCTCTGCTACTTTAAATACGTATTATTTT 24025
Qy 275 GlyGlySerIleGlyIleLysPheIleLeu 284
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24024 GGTGGGAGATTGGATTAGATTTCGGCTA 23995

RESULT 34
AF368000 831 bp DNA linear BCT 06-NOV-2001
LOCUS Cowdria ruminantium isolate Blaaukrans major antigenic protein MAP1
DEFINITION
gene, partial cds.
ACCESSION AF368000
VERSION AF368000.1 GI:15429315
KEYWORDS
SOURCE heartwater rickettsia.
ORGANISM Cowdria ruminantium
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE 1 (bases 1 to 831)
Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaïd,A.,
Haydon,D.T., van Heerden,H. and Allsopp,B.A.
Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants
Are Not Geographically Constrained and Show No Evidence of Having
Evolved under Positive Selection Pressure
J. Clin. Microbiol. 39 (11), 4200-4203 (2001)
JOURNAL 21539003
MEDLINE PUBMED 11682561
REFERENCE 2 (bases 1 to 831)
Allsopp,M.T., Dorfling,C., Maillard,J.-C., Bensaïd,A., Van
Heerden,H. and Allsopp,B.A.
Direct Submission
TITLE Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary
Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,
South Africa
FEATURES
source 1..831
/organism="Cowdria ruminantium"
/isolate="Blaaukrans"
/db_xref="taxon:779"
<1..>831
/codon_start=1
/transl_table=11
/product="major antigenic protein MAP1"
/protein_id="AAK98142.1"
/db_xref="GI:15429316"
```

```
/translation="ITSTLISLVFPGVSFSDVIEDSSPAGSVYISAKYMPATSHF
GKMSIKEDSKNTQVFLKQMDGVKVPYTSNNSTFTFKDYTSFRYNNPFLGFGAG
AIGYSMNGPRIEFVSVEFDVKPPGGNYKNDAMHMCALDQTAQTSNOGATLASSVW
KNEMLTDSLMLNACYDMLDMPSPYVCAGICTDLVSVINATNPKLISYOGKLGISY
SINEASIFIGGFHRRVIGNEFKDIATSKIFTSKASSINPFGASATLQVCHFGIE
I"
BASE COUNT 292 a 137 c 141 g 261 t
ORIGIN
Alignment Scores: 8.3e-26 Length: 831
Pred. NO.: 382.00 Matches: 101
Percent Similarity: 52.96% Conservative: 51
Best Local Similarity: 35.19% Mismatches: 93
Query Match: 25.30% Indels: 42
DB: 1 Gaps: 15
US-09-846-808-1 (1-284) x AF368000 (1-831)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16 ATATCATTTAGTGTCAATTTTACCTGGTGGTGGTCTTTCTGATGTAATACAGGAACACAGC 75
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaAargGlnHisLeuSer 48
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 AGCCCGCAGCGCAGTGTTTACATTAGCGCAAAATACATGCCCAACTGCATCATTTTGGT 135
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 49 LysLeuLeuIleLysGluSerAlaAlaasnThrValGluValPheGlyLeuLysLysasp 68
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 AAAATGTCAATCAAAAGAGATTCAAAAATACTCAACACAGTATTTGGTCTAAAAAAGAT 195
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsn----- 82
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 -----TGGGATGGCGTTAAAGTACCACACATCAGAAAACAGTAACAATTTCT 240
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 83 -----PheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 ACAATTTTACTGAAAAAGATATTCTTTTCAGATATGAAAAACAATCCGTTTGTAGTTT 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysasnPheArgIleGluSerGluLeuSerTyr 118
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GCTGGAGCAATTGGATCTCAATGAAT---GGGCCCAAGAATAGAGTTTGAAGTATCTAT 357
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 119 GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 GAAACTTTTGATGTCAAAAAATCCAGGTGCGCAACTATATAAAAT---GATGCACACATGTAC 414
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrVal 156
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 415 TGTGCTTTA-----GATACAGCACACGCCCAACTAGTAATCAAGGTGCA 456
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 157 ThrLeuIle-----AsnAsnGlyIleSerLeuThrSerAlaLeuIle 170
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 457 ACATGAGTTCATCTGTTATGTTAAGAAATTTTAAACAGATATATCATTAATATTTA 516
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 171 AsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIleThrTyrSerCys 188
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 517 AATGCATGTTATGATATATATGCTTGATGAATGCCA-----GTTCTCCCATATGTATGT 570
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 189 LeuGlyPheGlyValIaspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyr 208
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GCAGGTATTGGTACTGATTATTAGTGTCAAGTAATTAATGTACAAATCCTAAATATATCTTAT 630
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 209 GlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGly 228
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAAGAAAGCTAGGCATAGTACTCAATCTCAATCTCAAGTCTCTATCTTTATTTGGTGA 690
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 229 TyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrProCysasp 248
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 CATTTCCATAGATTATAGTAATGAATTTAAAGATATTGTACCTACCTCAAAA-----ATA 744
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 249 TyrProSerProThrProProAsnSerLysProHisValHisThrThrAlaLeuAlaMet 268
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

/organism="Cowdria ruminantium"
/isolate="Pokoase"
/db_xref="taxon:779"
<!.>816
/codon_start=1
/transl_table=11
/product="major antigenic protein MAP1"
/protein_id="AAK98153.1"
/db_xref="GI:15429338"
/translation="FITSLILSLVFLPGVSPFSDVIOENNPVGSVVISAKYMPYASH
FKMSTKEDSRDTKAVGLKQMDGVKTPSGNTSIFTEKDYFKYENNPFLGFAGAV
GYSNGPRIEFEVYTFDVPKPGNYKNDAMHCALDITASSTAGCATSVMYKNENL
TQISLMNACYDILMDGMPVPVPCAGIGTDLVSVINATNPKLSYQKGLGISVINPE
ASTIFGGHFRVIGNEFKDIATSKVPTSSGNASSAVSPGFASAILDVCHEGIEI"
BASE COUNT 283 a 129 c 148 g 256 t
ORIGIN

Alignment Scores:
Pred. No.: 9,01e-26 Length: 816
Score: 381.50 Matches: 93
Percent Similarity: 54.92% Conservative: 41
Best Local Similarity: 38.11% Mismatches: 83
Query Match: 25.26% Indels: 27
DB: 1 Gaps: 12

US-09-846-808-1 (1-284) x AF368013 (1-816)
QY 13 ValMetLeulleuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 19 ATATCAATTAGTGTATTTTACCTGGTGTCTCATTTTCTGATGTAATACAGGAAGAAC 78
QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 79 AATCCAGTAGTGTATTTATATAGCGCAAAATACATGCCAAGTGCATGCATTTTGGT 138
QY 49 LysLeuLeulleGlySerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 139 AAAATGTCATTAAGAAGATTTCTAGAGATACATAACACGATATTGGCTTAAAAAAGAT 198
QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsn-----Phe 83
Db 199 -----TGGGATGGTGTATTAACACACCAATCGGCTGATACACCAATTCATTTT 243
QY 84 AsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIle 101
Db 244 ACTGAAAAGACTATCTTTCAATATAGAAATAATCCATTTTAGTGTTCGAGGAGCA 303
QY 102 PheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPhe 121
Db 304 GTTGGATACCTCAATGAAT---GGACCAAGAATAAGATTTGAAGTATCTTATGAAACTTC 360
QY 122 HisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeu 139
Db 361 GACGTAAAAAATCCAGCGGTACTACAAAAAC-----GATGCACATATGTATGT 411
QY 140 AlaLysGluIleSerGlyLysSerAsnProAlaAsnAsnLysTyrValThrLeuIle 159
Db 412 GCTCTAGATACAGATCAACCTCTACTGCAGGAGCACTACATCT---GTTATGGTAAAA 468
QY 160 AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal-----Asp 177
Db 469 AATGAAATTTAAACAGATATTTTCATTAATGCTAAACGCATGTTATGACATAATGCTTGAC 528
QY 178 GlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp 197
Db 529 GGAATGCCA-----GTTTCTCCATATGTATGTGCAGGCATTTGCTAGCTAGTATCA 582
QY 198 PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr 217
Db 583 GTAATTAACGCTCAAAATCTTAATATCTATCAAGGAAAATAGGAATAAGTTATTC 642
QY 218 ValSerProGlnValSerValPheIleGluGlyTyrTyrThrHisGlyLeuPheGlyLys 237

```

```

Db 643 ATAAACCCGGAAGCTTCTATCTTTATTGGTGGGCACCTTCCATAGAGTCATAGGTAACGAA 702
QY 238 PheGluLysIle 241
Db 703 TTTAAGATATT 714
RESULT 37
AX042312 AX042312 861 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 8 from Patent WO0065063.
ACCESSION AX042312
VERSION AX042312.1 GI:11340972
KEYWORDS Ehrlichia chaffeensis.
SOURCE Ehrlichia chaffeensis.
ORGANISM Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE 1 (bases 1 to 861)
AUTHORS Barbet,A.F., Bowie,M.V., Ganta,R.R., Burridge,M.J., Mahan,S.M.,
Mcquire,T.C., Rurangirwa,F.R., Moreland,A.L., Simbi,B.H.,
Whitmire,W.W. and Alleman,A.R.
TITLE Nucleic acid vaccines against rickettsial diseases and methods of
use
JOURNAL Patent: WO 0065063-A 8 02-NOV-2000;
UNIVERSITY OF FLORIDA (US)
FEATURES
source
Location/Qualifiers
1..861
/organism="Ehrlichia chaffeensis"
/db_xref="taxon:945"
BASE COUNT 281 a 140 c 168 g 272 t
ORIGIN

Alignment Scores:
Pred. No.: 1.31e-25 Length: 861
Score: 380.00 Matches: 100
Percent Similarity: 48.67% Conservative: 46
Best Local Similarity: 33.33% Mismatches: 92
Query Match: 25.17% Indels: 62
DB: 6 Gaps: 14

US-09-846-808-1 (1-284) x AX042312 (1-861)
QY 14 MetLeulleuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn--- 32
Db 42 TTACTATGTCTTCTTACCTGGATATACCTTCTGTATCCAGTACAGGATGACACACATT 101
QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 102 AGTGGTAATTTCTACATCAGTGGAAAGTATATGCCAGCGCTTCGCAATTTGGAGTTTTT 161
QY 51 LeulleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp----- 68
Db 162 TCTGCCAAGGAAGAAATAACACAGTGTGGAGTATTTGGAATAGACCAAGATTGGGAT 221
QY 69 -----LeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThr 81
Db 222 AGATGTGTAATATCTAGAACCCACTTTAAGCGATATATTACC-----GTTCCA 269
QY 82 AsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPheSerGly 100
Db 270 AATTATTTCATTTAAG-----TATGAAATAATCTATTTTTCAGGATTTGCAGGA 317
QY 101 IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr 120
Db 318 GCTATTGGCTACTCAATGGAT---GGCCCAAGATAGAGCTTTGAAGTATCTTTATGAACA 374
QY 121 PheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHisPheAla 138
Db 375 TTCGATGTTAAAAATCAAGGTAACAATATATAAGAAC---GAAGCACATAGATATTATGCT 431
QY 139 LeuAla-----LysGluIleSerGlyGlySerAsnAsnProLysAsn 152
Db 432 CTGTCCCATCTCTTCGGCACAGACAGACATAGATGTT-----GCAGGC 476

```

```

Qy 153 AsnLysTyrValThrLeuIleAsnGlySerLeuThrSerAlaLeuIleAsnVal 172
   ::  ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 AGTGGCTGCTCTTCTAATAAATGAAGGACTACTGTGATAAATCATTTATGCTGAACGCA 536

Qy 173 CysTyrAspVal-----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGly 190
   ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 TGTATGATGATTAAGTAAGGACATACCTTTTCT-----CCTATATATATGTCAGGT 590

Qy 191 PheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGly 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 ATTGGTATTGATTAGTATCCATGTTTGAAGCTATAAATCTTAAATTTCTTATCAAGGA 650

Qy 211 LysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyr 230
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 651 AAATAGGCTTAAGTTACCTTATAGCCAGCAGGCTTCTGTGTTTATTTGGTGGACATTT 710

Qy 231 HisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrProCysAsp----- 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 711 CATAAGGTAGGAACGAATTTAGAGATATCTCTACTATGATACCTAGTGAATCAGCG 770

Qy 249 -----TyrProSerProThrProProAsnSerLysProHisValHis 262
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 771 CTTCGAGAAAGAACTACCTT----- 794

Qy 263 ThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlySerIleGlyIleLysPhe 282
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 -----GCAATAGTAGTAACACTGGACGTGTCTACTTTGGCATAGAACTTGGAGGAGGTTT 848

RESULT 38
LOCUS AF368011 834 bp DNA linear BCT 06-NOV-2001
DEFINITION Cowdria ruminantium isolate Nonile major antigenic protein MAP1
gene, partial cds.
ACCESSION AF368011
VERSION AF368011.1 GI:15429333
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichiae; Cowdria.
REFERENCE 1 (bases 1 to 834)
AUTHORS Allsopp,M.T., Dorfling,C., Maillard,J.C., Bensaid,A.,
Haydon,D.T., van Heerden,H. and Allsopp,B.A.
TITLE Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants
Are Not Geographically Constrained and Show No Evidence of Having
Evolved under Positive Selection Pressure
J. Clin. Microbiol. 39 (11), 4200-4203 (2001)
JOURNAL 21539003
MEDLINE 11682561
PUBMED
REFERENCE 2 (bases 1 to 834)
AUTHORS Allsopp,M.T., Dorfling,C., Maillard,J.C., Bensaid,A., Van
Heerden,H. and Allsopp,B.A.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary
Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,
South Africa

FEATURES
source 1..834
/organism="Cowdria ruminantium"
/isolate="Nonile"
/db_xref="taxon:779"
<1..>834
/codon_start=1
/transl_table=11
/product="major antigenic protein MAP1"
/protein_id="AAK98151.1"
/db_xref="GI:15429334"
/translations="FITSLISLVFLPGVSFSDVIOEDSNPAGSVISAKYMPFASH
FGKMSIKEDSKNTQVFLKKDGVKVPSTENTYSSLETKXDYSRYENNEPLGFA
CATGYSMNGPRIFEVSYETFDVKNPGGNKNDAHMYCALDQAQSNATGATLASSVM
IKNENLTNLSIMLNACDYLMDLGDGKVPSPYVACIGDGLDLSVINATNPKLSYQKGLGIS
YSINSEASIFIGGHFHVIRVIGNEFKDIATLKIFTATNKNVSTVNPFGFASATLDVCHFGI"

CDS

```

```

BASE COUNT 250 a 133 c 143 g 268 t
ORIGIN
Alignment Scores: 1.4e-25 Length: 834
Pred. No.: 379.50 Matches: 91
Score: 53.04% Conservative: 40
Percent Similarity: 36.84% Mismatches: 89
Best Local Similarity: 25.13% Indels: 27
Query Match: 1 Gaps: 10
DB:
US-09-846-808-1 (1-284) x AF368011 (1-834)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
   ::  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 ATATCATTTAGTCTCATTTTACCTGGTGTGCTCTTTCTGTGATTAATACAGGAGGATACG 78

Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 AACCAGCTGGTAGTGTTTACATTAGCGCAAAATACATGCCCAACTGCTTCACACTTTGGT 138

Qy 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 AAAATGTCAAATTAAGAAGATTCAAAAATACTCAACAGTGTTTGGTCTAAAAAAAGAT 198

Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPheAsn 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 -----TGGGATGGAGCTTAAAGTACCACATCAGAAAACCAATTTACTCT 243

Qy 85 Ile-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 TCACCTTTTACTGAAAAAGATTATTCTTTCAGATATGAAAAACAATCCGTTTGTAGGTTT 303

Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerCyluLeuSerTyr 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 GCTGGCAATTTGGGTACTCAATGAAT---GGACCAAGAATAGATGTCGAAGTATCCCTAT 360

Qy 119 GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GAACTTTTGATGTAAAAACCCAGGTGCGCAACTACAAAAAT--GATCCACACATGTATC 417

Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrVal 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 TGTGCTTTAGATACAGCAACAGAGTGTCTACTAATGTGTGCAACATTAGCTTTCATCTGTT 477

Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 ATGATAAAAAATGAAAAATTTAACAAAATATATCAATTAATGTTAAATGCCGTGTTATGATA 537

Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 ATGCTTGATGGA-----AAGCCAGTTTCTCCATATGATGTGTCAGGATTTGCTACTGAT 591

Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 TTAGTGTCAAGTAATTAATGCTACAAATCCTAAATTTATCTTATCAAGGAAAGCTAGGCATA 651

Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPhe 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 AGTTACTCAATCAATTCGAAGCTTCTATCTTTATCGGTGGACATTTCCATAGAGTTATA 711

Qy 235 GlyLysLysPheGluLysIle 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 GGTAAATGAATTAAGATATT 732

RESULT 39
LOCUS AF368008 825 bp DNA linear BCT 06-NOV-2001
DEFINITION Cowdria ruminantium isolate Wara87/7 major antigenic protein MAP1
gene, partial cds.
ACCESSION AF368008
VERSION AF368008.1 GI:15429327

```


KEYWORDS

SOURCE heartwater rickettsia.

ORGANISM

Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichiae; Cowdria.

REFERENCE

1 (bases 1 to 825)

AUTHORS

Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaïd,A.,

Haydon,D.T., van Heerden,H. and Allsopp,B.A.

Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants

Are Not Geographically Constrained and Show No Evidence of Having

Evolved under Positive Selection Pressure

J. Clin. Microbiol. 39 (11), 4200-4203 (2001)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

Direct Submission

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

Qy 119 GluThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHis 136

Db 361 GAAACTTTTGTATGTAATAAAACCCAGGCGCAACTATAAAAC---AACGCACACATGTAC 417

Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156

Db 418 TGTCTTTAGTATACAGCAGCAGCAAAATAGCAGTATGCGCAGGATTAACTACATCTGTT 477

Qy 157 ThrLeuLeuAsnGlyIleSerLeuThrSerAlaLeuLeuAsnValCysTyrAspVal 176

Db 478 ATGGTAAAAACGAAATTTAACAAATATATCATTAATGTTAAATGCGGTATGATATC 537

Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194

Db 538 ATGCTTGTGAATACCA-----GTTTCTCCATATGATGTGCAGGTATGGCACTGAC 591

Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214

Db 592 TTAGTGTGAGTAATTAATGCTACAAATCCTAAATATCTTATCAAGGAAGCTAGGCATA 651

Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234

Db 652 AGTTACTCAATCAATCTCAAGCTTCTATCTTATCGGTGGACATTTCCATAGAGTTATA 711

Qy 235 GlyLysLysPheGluLysIle 241

Db 712 GGTAAATGAATTTAAAGATATT 732

RESULT 40

AF368010

LOCUS

DEFINITION

Cowdria ruminantium isolate Morgenswag2 major antigenic protein

MAP1 gene, partial cds.

ACCESSION

AF368010

VERSION

AF368010.1 GI:15429331

KEYWORDS

SOURCE

heartwater rickettsia.

ORGANISM

Cowdria ruminantium

Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

Rickettsiaceae; Ehrlichiae; Cowdria.

REFERENCE

1 (bases 1 to 834)

AUTHORS

Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaïd,A.,

Haydon,D.T., van Heerden,H. and Allsopp,B.A.

Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants

Are Not Geographically Constrained and Show No Evidence of Having

Evolved under Positive Selection Pressure

J. Clin. Microbiol. 39 (11), 4200-4203 (2001)

JOURNAL

MEDLINE

PUBMED

REFERENCE

2 (bases 1 to 834)

AUTHORS

Allsopp,M.T., Dorfling,C., Maillard,J.-C., Bensaïd,A., Van

Heerden,H. and Allsopp,B.A.

Direct Submission

TITLE

Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary

Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,

South Africa

JOURNAL

TITLE

Location/Qualifiers

1. .834

/organism="Cowdria ruminantium"

/isolate="Morgenswag2"

/specific_host="goat"

/db_xref="taxon:779"

/note="detected in goat blood from the farm Morgenswag,

Northern Cape, South Africa"

<1. .>834

/codon_start=1

/transl_table=11

/product="major antigenic protein MAP1"

/protein_id="AAK98150.1"

/db_xref="GI:15429332"

/translation="FITSLISLVFPLGVSFSDVIOEDSNPAGSVVISAKYMPGTASH

FGKMSKEDSKNTQTVFGLKDWGVKPTSENTNYSLSLFTKDYSPRYENNPFLGFA

CDS

heartwater rickettsia.

Cowdria ruminantium

Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichiae; Cowdria.

1 (bases 1 to 825)

Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaïd,A.,

Haydon,D.T., van Heerden,H. and Allsopp,B.A.

Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants

Are Not Geographically Constrained and Show No Evidence of Having

Evolved under Positive Selection Pressure

J. Clin. Microbiol. 39 (11), 4200-4203 (2001)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Allsopp,M.T., Dorfling,C., Maillard,J.-C., Bensaïd,A., Van

Heerden,H. and Allsopp,B.A.

Direct Submission

TITLE

Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary

Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,

South Africa

JOURNAL

TITLE

Location/Qualifiers

1. .825

/organism="Cowdria ruminantium"

/isolate="Mara87/7"

/db_xref="taxon:779"

<1. .>825

/codon_start=1

/transl_table=11

/product="major antigenic protein MAP1"

/protein_id="AAK98148.1"

/db_xref="GI:15429328"

/translation="FITSLISLVFPLGVSFSDVIOEDSNPAGSVVISAKYMPGTASH

FGKMSKEDSKNTQTVFGLKDWGVKPTSENTNYSLSLFTKDYSPRYENNPFLGFA

GAIGYSMNGPRFEFYSYETFDVKNPGGKNNHMYCALDRAQNSTGACGLTTSVY

VNNENITSLMLNAGTIDMLGIPVPCAGIGDLVSVINATPKLSYQKLGIS

YSINSASISFIGHFRHVGNEFKDITATIKFTSKTGISNPGFASATLDVCHRGIEI"

BASE COUNT

293 a 139 c 138 g 255 t

ORIGIN

Alignment Scores:

Pred. No.: 1.71e-25 Length: 825

Score: 378.50 Matches: 91

Percent Similarity: 52.63% Conservative: 39

Best Local Similarity: 36.84% Mismatches: 90

Query Match: 25.07% Indels: 27

DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF368008 (1-825)

Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28

Db 19 ATATCATAGTGTGCTTTTACCTGGTGTCTCTTTCTGATGTAATACAGGAAGACAGC 78

Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48

Db 79 AACCCAGCAGCAGTGTTCATTAGGCGCAAAATACATGCCAAGTGTGATGCTGATTTTGGT 138

Qy 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68

Db 139 AAATGTCAATCAAGAAGATGTAATAAAATACCTCAAAACAGTATTTTGTCTAAAAAAGAT 198

Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAspPhe 83

Db 199 -----TGGGATGGGTAAACACACCATCATGATCTAGCATATAATCT 243

Qy 84 AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg----LeuGlyPhe 98

Db 244 ACAATTTTACTAAAAAGACTATCTTTCAGATATGAAACAAATCGGTTTATAGGTTTC 303

Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118

Db 304 GCTGGAGCAATTGGGTACTCAATGAAT---GGACCAAGAATAGAGTTTCGAAGTATCCTAT 360

GAIGYSMNGPREEVSYETFDVKNPGGNYKNDAMHYCALDTAQSATNGATLASSVV
IKNENLTNISMLNACDYIMLDGMPVSPYVCAGIGPDLVSVINATNPKLSYOGKLGIS
YSINSGASIFIGCHFHRVIGNEFKDIATKIFTATNKVSTVANPGFSATLDDVCHFGI
EI"

BASE COUNT 285 a 136 c 145 g 268 t
ORIGIN

Alignment Scores:

Pred. No.: 1.73e-25 Length: 834
Score: 378.50 Matches: 91
Percent Similarity: 53.04% Conservative: 40
Best Local Similarity: 36.84% Mismatches: 89
Query Match: 25.07% Indels: 27
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF368010 (1-834)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
DB 19 ATTTCAATAGTGTCTATTTTACCTGGCGTGTCTCTTCTGATGTAATACAGGAGGATAGC 78
QY 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
DB 79 ACCCAGCTGGTAGTGTTCATATAGCGCAAAATACATGCCAACCTCTTCACACITTTGGT 138
QY 49 LysLeuLeuIleLysGluSerAlaAlaThrValGluValPheGlyLeuLysLysAsp 68
DB 139 AAAATGTCATTAAGAAGATTCAAAAATACTCAAAACAGTGTGGTCTAAAAAAGAT 198
QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPheAsn 84
DB 199 -----TGGGATGGAGTTAAAGTAGCACCAACATCACAAACACCAATTAATCTCT 243
QY 85 Ile-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
DB 244 TCACITTTTACTGAAAAGATTATCTTTCAGATATGAAAACAATCCGTTTGTAGGTTT 303
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
DB 304 GCTGGAGCAATTGGGTACTCAATGAAT---GGACCAAGAATAGAGTTCCAGATATCCTAT 360
QY 119 GluThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHis 136
DB 361 GAACTTTTGTATGTAACCAACCCAGGTGGCAGTACAAAAAT---GATGCACACATGTAC 417
QY 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
DB 418 TGTGCTTTAGATACAGCACACAGAGTGTCTAATGTTGGCAACATTAGCTTCATCTGTT 477
QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystyrAspVal 176
DB 478 GTGATAAAAAATGAAAAATTAACAATAATATACATAATGTTAAATGCGTGTATGATATA 537
QY 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValasp 194
DB 538 ATGCTGTAGTGAATGCCA-----GTTCTCCATATGTATGTCAGGTATTGGTCTGAT 591
QY 195 ThrIleAspPheLeuSerLysTyrThrLysPheSerTyrGlnGlyLeuGlyAla 214
DB 592 TTAGTGTCAAGTAATTAATGCTACAAATCTCAATATCTTATCAAGGAAGTAGGCATA 651
QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
DB 652 AGTACTCAATCAATCTCGAGTCTCTATCTTTATCGGTGGACATTTCATAGAGTTATA 711
QY 235 GlyLysLysPheGluLysIle 241
DB 712 GGTAAATGAATTTAAAGATATT 732

Search completed: October 19, 2002, 04:29:56
Job time : 1921 secs

This Page Blank (uspto)

ORGANISM Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE 1 (bases 1 to 14759)
Yu.X.J., McBride,J.W., Zhang,X. and Walker,D.H.
AUTHORS Characterization of the complete transcriptionally active Ehrlichia
TITLE Characterization 28 kDa outer membrane protein multigene family
JOURNAL chaffeensis 28 kDa outer membrane protein multigene family
VOLUME 248 (1-2), 29-68 (2000)
PAGES 29-68
AUTHORS 2 (bases 1 to 14759)
Yu.X.J., McBride,J.W., Zhang,X. and Walker,D.H.
AUTHORS Direct Submission
JOURNAL Submitted (01-FEB-2000) Pathology, University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA
FEATURES
source
1..14759
/organism="Ehrlichia chaffeensis"
/strain="Arkansas"
/db_xref="taxon:945"
400..1731
/gene="clpX"
400..1731
/gene="clpX"
/note="Arp-dependent protease"
/codon_start=1
/transl_table=11
/product="ClpX"
/protein_id="AAF73407.1"
/db_xref="GI:8132824"
/translation="MADNEKNSCSFCGKIHSEVRKLIAGPSVFTCNEDICLSCGIL
QEESYKTDPLKPKKEIKKVLDEYVIGQESKKKVLISVAVNHYKRLSNLSVSE
EISKNSVLLIGPTGSGKTLARTARVLIQIPAMADATLTLEAGVGDENILKLL
QAANFNAAQGGIYIDDEVKISRKSENTITRDVSGEGVOQALLKLVIEGTVSVPV
QGRKHPOASHLSKSVLMQLSNRKVSSTTKVTALYITETIQDFEQLSDQEEV
LELNYEGATRDVDKQYKEDNRYFALCKKDS TESTDNGSNHVMKNEGFRVLSL
TNACYDMIVSNSSLPVACIGIGOTIFNLGNTNHTLEKAKLGLGLSPKTVIFA
NGYIVKTKDNTSNLSVQIPLELKEAPRHIDPIACFNADNYGGEVGLRFL"
2236..3051
/gene="p28-1"
2236..3051
/gene="p28-1"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-1"
/protein_id="AAF73408.1"
/db_xref="GI:8132825"
/translation="MLILTFPHLSLAVLDHNSVMYVGIQKPAROHLKLLIKESAA
NTEVEGLKDLNLLDTGKONTNFIKNYENNNPANNRIVYLINNGISLTALINVC
LSYETFIKNNGYKRIKCEKHFALEISGGSNPANNRIVYLINNGISLTALINVC
YVDVGLKHLNITVSCIGFGVDITDFLSKTYTKFSGQKLGASITVSPQSVFIEGYH
GLFGKAFKIPVNYPCDYSPTPPNSKPHVHTALAMLSIGYGGSIGIKFIL"
3158..3604
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAF73409.1"
/db_xref="GI:8132828"
/translation="MNNKSRGQVLLFAAYSIFKYINCSCTKFRSLVKCFVYNPLAYF
YKSTFLKGLKFFVNYKTSNLSHRSLKFLGIIRTNFKISLNFNPNRVLIY
YKILNKCNSNINLLNKYTNPTATSFINSVIIGIYGFVYNNY"
3614..4507
/gene="p28-2"
3614..4507
/gene="p28-2"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-2"
/protein_id="AAF73410.1"
/db_xref="GI:8132831"
/translation="MSKKKFTITGTVLASLLSFLSIESFSAINHNHTGNNTSGIYITG
QIRPGYSHFSNFSVKETVDYTIOLGVYKKSASSIDPNTYSNFGQPYTVTFQDAAFS
GAIGYSIPESLRLELEGSKEKDPKDYSAKDAREFALARTSTTVPDQAKYTM
KNGLSVASIMINGCYDLFSNNLVSPYICAGGEDFIEFDTFLHKLAYQCKLGISY
YFFPKINVFAGGYHRRVNNVHTVPDEFPKATFSAVATLUNVAFGEAGCV"
/db_xref="GI:8132832"
/translation="MSYAKVFLICLILLVPSLSFAIVNNDLKNIGHFYIGGYKPK
GYPRNRELVTNNIRELMSDEECRTIPIHMQVSAQGTLPPEALBELADKFPBEGY
LYETIPNPTNNKLLAGGVIGYSTHERVEAFYDFKNTLAPAGYLHKNFYETFA
LATMDTKRHQSAEDKYYMKNTGITLSPFIINACYDILTKTRNAVYPLCLGVGGN
FIDFLDQSEKFAQYQAKVGISYFVSPNIAFFIDGSPFHGLNNQFSDSPVVDYSSGGP
TISAKFNAPLASSIGIRIS"
4524..5381
/gene="p28-3"
4524..5381
/gene="p28-3"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-3"
/protein_id="AAF73411.1"
/db_xref="GI:8132838"
/translation="MOKLYISFILLGSLLPKVFVFCMHQNNIDGSYVTIKYQLTTPH
FKNFIKETDFDQEPIGLAKITANTKFDTLKENFSPHLQTDSDSYQNDLLGIGL
SVGLFVKSPRIEPEGAYKNFTKELARYKSKDGYKYFAIPKSEHGFNDFTGVTVAK
NNGIISNIINLCSETKYSKFTYICIGYVGGDFIEFDYMRLEKFAQGVGYSPYT
SKILTSINGQYHKYVIGNKFEPLPVYQPVLEKRLVNTKTSKDIDQDVTASLTNLNLEHS
SEIGLSPIF"
5394..6212
/gene="p28-4"
5394..6212
/gene="p28-4"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-4"
/protein_id="AAF73412.1"
/db_xref="GI:8132829"
/translation="MYMYNKKHYCIYVTVTLFPLLPISLSALICNGVEKDLKVS
TVSSQYKPSIFHFNFESIQESHPPKSEEFKIKANLNILKNAYNLQDQNTSF
SGTIGYFSKGLLEAEGCYQEFNVKNSNLSIISNKHYSRIHDENYAITNNKLUSA
SIMVNTCYDISINNTSIVPYLCTGIGEDVLGTFNTIHFKLAYQKVGMSYLINNILL
ESDIYHYKVMGNRFKNLYQYVADPNISEETIPIILAKLDIGYFSGEIGRPFMFN"
6228..7115
/gene="p28-5"
6228..7115
/gene="p28-5"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-5"
/protein_id="AAF73413.1"
/db_xref="GI:8132830"
/translation="NTKKNFVNVILTFLLPLPKSFSTTYANNNTITQKVLVYISGG
YKPSIPHKNSFVENDKVDLIGLTVDTVYITEHLRDNTEKFNTHYIAKKNKNIIF
SSAIGYSGQPRLEIESYGDVNVNKNYAVQDNYRFALVRKNGSNFSPKPHET
SQPSDSNPKKSYITLMKNGNVEVASVINGCYDFSNNTTISPYVICIGVGGDFIEFFE
VMHIFACQKSGVSIYSPISPTITFADAHYHKVINNNKNNLHVKSYSLEKNSPTITISA
TAKLANIEYFGGVMGRFIF"
7137..7976
/gene="p28-6"
7137..7976
/gene="p28-6"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-6"
/protein_id="AAF73414.1"
/db_xref="GI:8132831"
/translation="MSKKKFTITGTVLASLLSFLSIESFSAINHNHTGNNTSGIYITG
QIRPGYSHFSNFSVKETVDYTIOLGVYKKSASSIDPNTYSNFGQPYTVTFQDAAFS
GAIGYSIPESLRLELEGSKEKDPKDYSAKDAREFALARTSTTVPDQAKYTM
KNGLSVASIMINGCYDLFSNNLVSPYICAGGEDFIEFDTFLHKLAYQCKLGISY
YFFPKINVFAGGYHRRVNNVHTVPDEFPKATFSAVATLUNVAFGEAGCV"

This Page Blank (uspto)

CDS
 8214..8591
 /note="orf378"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAF73415.1"
 /db_xref="GI:8132832"
 /translation="MNCYLKPKLNIQIFYGIFSRNFQNTLLIIPNDSKCGFYTL
 WNPALHYTVTLGSSYRNPFDFILYENIIOCKLLINRSVNLQHNKNTLVIPIPN
 AREFSNEIRVRNISINKSESYEC"
 8581..9432
 /gene="p28-7"
 8581..9432
 /gene="p28-7"
 /note="component of 28-kDa outer membrane protein
 multigene family"
 /codon_start=1
 /transl_table=11
 /product="p28-7"
 /protein_id="AAF73416.1"
 /db_xref="GI:8132833"
 /translation="MSAKKLFITIGSVLVCLVSLYPTKSLSLNNLNNTKCTGLVYS
 GQKPVSHFSNFKETVTDKELGLAKDKISDITTNKFNPIPIYKFKODNAVS
 FSAVGIISODSPREVEWSYEEDVKNFNGVYVSEAFYIALARGIDNQLKYPETNK
 YVVKNGLSVASIIICGYDFSLNKLKVPYICVGGDIIIEFFSAVSFKAYQGV
 GSYPLFSNMIIIFADGYHKVIGNKNLNLVQHVSLNSHPKSTFAVATLNVEYFGE
 FGLKFTF"
 9441..10268
 /gene="p28-8"
 9441..10268
 /gene="p28-8"
 /note="component of 28-kDa outer membrane protein
 multigene family"
 /codon_start=1
 /transl_table=11
 /product="p28-8"
 /protein_id="AAF73417.1"
 /db_xref="GI:8132834"
 /translation="MSKKNFITIGATLIHMLNPISFPETINNNTDKLSGLVSGQYK
 PGISHPSKSVKELYNDIOLIGLHNAISTLINLNTDFNPKYKFFONNTISQYK
 IGYSDTGARFELGSEYEDVDPGDLIKDTIRYFALARNPSSSPSNNTVWRN
 DGVSITFIKNGYDFIKDLKLEVPYCVGVGGDFIEFFDALHKLAYQCKLGINYL
 STQASVFDIGYHKVIGNKNLNLVQHVASTDFGPVYAVATLNIGYFGEIGRLTF"
 10278..11135
 /gene="p28-9"
 10278..11135

Alignment Scores:

Pred. No.: 2,6e-126 Length: 14759
 Score: 1510.00 Matches: 284
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-846-808-1 (1-284) x AF230642 (1-14759)

QY 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuLeuPheThrPro 20
 Db 2197 ATGAGTAAAGGAGTACACAAAGTTGTTTGTGGTGTATGTTAATATTATTACACCA 2256
 QY 21 HisLeuSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValGlyLeuGlnTyr 40
 Db 2257 CATATTCTTTAGCATCAGTTTAAATGACCAATAATCTATGATGTTGTGTTTACGTAC 2316
 QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaAlaAsnThrVal 60
 Db 2317 AAGCCAGCCAGACCAACCTTATCAAAACTCTTATAAAGAAAGCGACGCAACACTGTA 2376
 QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyLysAspAsn 80
 Db 2377 GAAGTTTTCGGGTAAAGAAAGACTTACTAAATGATCTATTAACTGGTATTAAAGATAAT 2436

QY 81 ThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly 100
 Db 2437 ACTAATTTTAAATAAATACACCCATATATGAAATAATATCGTTTAGGTTTTTCGGGT 2496
 QY 101 IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr 120
 Db 2497 ATATTGGATATTACTATAACAAAAATTCAGATAAATCTGAACCTTCTTACGAAACT 2556
 QY 121 PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
 Db 2557 TTCACATAAAGAACACACGGATATAAAGAAATGATTGTGAAAGCAATTTTGCCTTTGCA 2616
 QY 141 LysGluIleSerGlyGlySerAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
 Db 2617 AAGGAATAAGCGGTGTAGTACAAATCTCGCCATATAAATATGTTACCTTAATAAT 2676
 QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValaspGlyLeuLys 180
 Db 2677 AATGGTATTTCACTTACCTCAGCTCTGATCAAGCTTTGTTATGATGTTGACGGACTTAA 2736
 QY 181 HisAsnIleIleThrTyrSerCysLeuGlyPheGlyValaspThrIleAspPheLeuSer 200
 Db 2737 CATATAATAATAACGTATTTCATGTTTAGGATTTGGAGTAGACACAAATGATTTTCTCAGT 2796
 QY 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
 Db 2797 AAGTATACTACAAGTTTTCATACCAAGGAAATTAGAGCTAGTTATCTGTTCTCCT 2856
 QY 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys 240
 Db 2857 CAAGTATCGGTTTTTATAGAAGCTTATTATCATGTTTATTGTTAAATAATTTGACAAA 2916
 QY 241 IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis 260
 Db 2917 ATTCTCTGTTAAATTTATCCATGCTACTATCCATCCACCAACACCAACAGCAAGCCACAT 2976
 QY 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyLe 280
 Db 2977 GTTCACATACACATAGTATGTTAAGTATTGGATATTACGGTGGGAGTATTGGANTA 3036
 QY 281 LysPheIleLeu 284
 Db 3037 AAGTTTATATTA 3048
 RESULT 2
 ECUT2291
 LOCUS
 DEFINITION
 ECUT2291
 Ehrlichia chaffeensis strain Arkansas major outer membrane protein
 Amp-1 multigene cluster, complete sequence.
 U72291 AF021338
 U72291.2 GI:13511827
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Ehrlichia chaffeensis.
 Ehrlichia chaffeensis
 Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
 1 (bases 1484 to 21136; 21479 to 22234).
 Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.
 Immunodominant major outer membrane proteins of Ehrlichia
 chaffeensis are encoded by a polymorphic multigene family
 Infect. Immun. 66 (1), 132-139 (1998)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 2 (bases 1 to 27190)
 Ohashi,N., Rikihisa,Y. and Unver,A.
 Analysis of Transcriptionally Active Gene Clusters of Major Outer
 Membrane Protein Multigene Family in Ehrlichia canis and E.
 chaffeensis
 Infect. Immun. 69 (4), 2083-2091 (2001)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 3 (bases 1 to 27190)

This Page Blank (uspto)

product-market

product-market

This Page Blank (uspto)

/product="major outer membrane protein OMP-1V"
/protein_id="AAK28665.1"
/db_xref="GI:13511836"
/translation="MSKKRTITGTVLALLSLFSLIESFSAINHNHTGNTSGIYITG
QYRGVSHFNFSVKTNDVTIQLGVYKKKSSSIDPNTYINFGQPYTVPFDNAASF
GAIGYSPESLRUELEGSYEKFDKPKYSAKDAFFALARNSTVTPDAQKYTVM
KNGLSVASIMINGCYDLSPNNLVSPYICAGIGEDIEFFDLHIKLAYOGKLGISY
YFFPKINVFAGGYHVRVIGNKFNLVNHNVTLDFFPKATSAVATLVAVFGGEGVK
FTF"
8300..8890
/codon_start=1
/transl_table=11
/product="unknown function UN3"
/protein_id="AAK28666.1"
/db_xref="GI:13511837"
/translation="MIYKELTRVGEVILAYLSFILSTYIFLVLNIRVNSLAICVI
SLUKTNFVSTKLKDKCRDYSNNCYLYGKPLNQIFYGIFSFIRNFQNTLI
IPNDSKGVFTLWNPALHYTTLTGSEYRNFEDILYENIIOCKKILLINRSVLNQ
HNKNTLVIIPIPNAREFSNIRVNRINISINKSEVYC"
8880..9731
/gene="omp-1w"
8880..9731
/gene="omp-1w"
/note="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1W"
/protein_id="AAK28667.1"
Alignment Scores:
Pred. No.: 5,45e-126 Length: 27190
Score: 1510.00 Matches: 284
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-846-808-1 (1-284) x ECU72291 (1-27190)
Qy 1 MetSerLysArgSerAsnArgLysPheValLeuThrPValMetLeuLeuLeuLeuPheThrPro 20
Db 2498 ATGAGTAAAGAGGTACAGAAAGTTGTTTGGGGTTATGTTAAATATTATTACACCA 2557
Qy 21 HistSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValGlyIleGlnTyr 40
Db 2558 CATATTTCTTAGCATCAGTTTAAATGACCAATATCTATGATGTTGGTATTACGTAC 2617
Qy 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaAlaAsnThrVal 60
Db 2618 AAGCCAGCCAGACACACCTATCAAACTCTTTATAAAGAAAGCGCAGCAACACGTGA 2677
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 2678 GAAGTTTTGGGTTAAAAAAGAGCTTACTAAATGATCATTAATGTTGTTAAAGATAAT 2737
Qy 81 ThrAsnPheAsnLysLysAspProTyrTyrGluAsnAsnArgLeuGlyPheSerGly 100
Db 2738 ACTAATTTTAAATAAATACACCATATATGAAATATATCGTTAGGTTTTCGGGT 2797
Qy 101 IlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr 120
Db 2798 ATATTTGGATATTACTATACAAAATTTTCAGATAGAACTGAATCTGACTTCTTACGAAC 2857
Qy 121 PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
Db 2858 TTCACATAAAGACACACGATATAAAGAAATGATTGTGAAAAGCAATTTGCTCTTGCA 2917
Qy 141 LysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
Db 2918 AAGAAATAAGCGGTGGTAGTACAAATCCCGCAATAAATAATATGTTACCTTAATAAT 2977
Qy 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystTyrAspValAspGlyLeuLys 180

Db 2978 AATGGTATTTTCACTTACCTCAGCTCTGATCAACGTTGTTTATGATGTTGACGGACTTAA 3037
Qy 181 HisAsnIleLeuThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
Db 3038 CATAATAATAACGATATTCATGTTTAGGATTTGAGTAGACACAATTCATTTCTCAGT 3097
Qy 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
Db 3098 AAGTATCTACAAAGTTTTCATACCAAGGAAATTAGAGCTAGTATTACTGTTTCTCCT 3157
Qy 221 GlnValSerValPheIleGluGlyTyrThrHisGlyLeuPheGlyLysLysPheGluLys 240
Db 3158 CAAGTATCGGTTTTTATAGAAGTTATATCATGCTGTTTATTTGGTAAAAAATTTGAGAA 3217
Qy 241 IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis 260
Db 3218 ATTCCTGTAATATTCATGATGACTATCCATCACCACCAACCAACAGCAAGCCACAT 3277
Qy 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIle 280
Db 3278 GTTCACACTACAGCATAGCTATGTTAAGTATTGGATATTACGGTGGGAGTATTGGAATA 3337
Qy 281 LysPheIleLeu 284
Db 3338 AAGTTTATATA 3349
RESULT 3
AF078553
LOCUS
DEFINITION 28254 bp DNA linear BCT 02-APR-2001
Ehrlichia canis major outer membrane protein P30 multigene cluster
1. complete sequence
ACCESSION AF078553 AF078554 AF078555 AH006958
VERSION AF078553.2 GI:13512584
KEYWORDS
SOURCE Ehrlichia canis.
ORGANISM Ehrlichia canis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
AUTHORS 1 (bases 1 to 28254)
TITLE Cloning and characterization of multigenes encoding the
Immunodominant 30-kilodalton major outer membrane proteins of
Ehrlichia canis and application of the recombinant protein for
serodiagnosis
J. Clin. Microbiol. 36 (9), 2671-2680 (1998)
98371112
REFERENCE 2 (bases 1 to 28254)
PUBMED 9705412
AUTHORS Ohashi, N., Unver, A., Zhi, N. and Rikihisa, Y.
TITLE Analysis of transcriptionally Active Gene Clusters of Major Outer
Membrane Protein Multigene Family in Ehrlichia canis and E.
chaffeensis
Infect. Immun. 69 (4), 2083-2091 (2001)
21153566
REFERENCE 3 (bases 1 to 28254)
PUBMED 11254561
AUTHORS Ohashi, N., Unver, A., Zhi, N. and Rikihisa, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1998) Department of Veterinary Biosciences, The
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA
REFERENCE 4 (bases 1 to 28254)
PUBMED 11254561
AUTHORS Ohashi, N., Rikihisa, Y. and Unver, A.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2000) Department of Veterinary Biosciences, The
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA
REMARK Sequence update by submitter
COMMENT On Apr 2, 2001 this sequence version replaced gi:3790556 gi:3790555
gi:3790558 gi:3790557.
FEATURES
source Location/Qualifiers
1..28254
/organism="Ehrlichia canis"
/strain="Oklahoma"
/db_xref="taxon:944"

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)